

Random walks in Euclidean space

Péter Pál Varjú*

March 4, 2013

Abstract

Fix a probability measure on the space of isometries of Euclidean space \mathbf{R}^d . Let $Y_0 = 0, Y_1, Y_2, \dots \in \mathbf{R}^d$ be a sequence of random points such that Y_{l+1} is the image of Y_l under a random isometry of the previously fixed probability law, which is independent of Y_l . We prove a local limit theorem for Y_l under necessary non-degeneracy conditions. Moreover, under more restrictive but still general conditions we give a quantitative estimate which describes the behavior of the law of Y_l on scales $e^{-cl^{1/4}} < r < l^{1/2}$.

1 Introduction

Let X_1, X_2, \dots be independent identically distributed random isometries of Euclidean space \mathbf{R}^d . Let $x_0 \in \mathbf{R}^d$ be any point and consider the sequence of points

$$Y_0 = x_0, \dots, Y_l = X_l(X_{l-1}(\dots(x_0))), \dots$$

We call these the random walk started from the point x_0 , and Y_l is its l th step.

The purpose of this paper is to understand the distribution of Y_l .

This problem can be traced back to Arnold and Krylov [2] who studied the mixing of the random walk on the sphere where the steps are rotations. They asked if their results extend to isometries of Euclidean or hyperbolic space.

*I acknowledge the support of the European Research Council (Advanced Research Grant 267259)

Existing results in the literature can be divided into two classes. Some papers describe the behavior of the measure on scale $O(1)$ others do it on scale $O(\sqrt{l})$. We begin by discussing the first category.

Každan [16] and Guivarc'h [13] proved a ratio limit theorem for $d = 2$. This result describes the local behavior of the distribution of Y_l . It states that the conditional distribution of Y_l on a fixed compact set is asymptotically uniform, i.e. Lebesgue. More precisely, for any two smooth compactly supported functions f and g we have

$$\lim_{l \rightarrow \infty} \frac{\mathbf{E}[f(Y_l)]}{\mathbf{E}[g(Y_l)]} = \frac{\int f(x)dx}{\int g(x)dx}$$

provided that the denominator on the right do not vanish. The distribution of X_1 of course need to satisfy some natural non-degeneracy conditions for which we refer the reader to the original articles. The proofs rely on the fact that $\text{SO}(2)$ is commutative.

In the papers [3], [17], [20] the local limit theorem is generalized to higher dimension, however the arguments require some restrictive assumption on the law of X_1 , e.g. absolute continuity, which implies that the group generated by the support of X_1 contains translations. In the absence of translations new ideas are required to obtain the local limit theorem in full generality which is the main goal of our paper.

Very recently, Conze and Guivarc'h [9] proved a ratio limit theorem under certain assumption on the associated random walk on $\text{SO}(d)$. This assumption may hold in full generality but it has been verified only under special circumstances yet. (We elaborate on this assumption in Section 3 after Theorem A.) Their approach also does not rely on translations, but differ from the methods of this paper.

Tutubalin [23] proved a central limit theorem for dimension $d = 2$ and $d = 3$, which was later generalized to higher dimension by Gorostiza [14] and Roynette [22]. The central limit theorem describes the behavior of the distribution of Y_l on scale $O(\sqrt{l})$. More precisely, it claims that Y_l/\sqrt{l} converges weakly to a Gaussian distribution if Y_1 has finite second moments. The central limit theorem was revisited by many authors, see e.g. [15], [21], [18] and [1]. In these works the central limit theorem was even generalized to cases when the distribution Y_1 does not have second finite moments and the limit distribution is not Gaussian.

To formulate our results we need to make some non-degeneracy condition on the law of X_i . We say that X_i are degenerate if there is a proper closed

subset A of \mathbf{R}^d and an isometry $\gamma \in \mathbf{R}^d$ such that Y_l is almost surely contained in $\gamma^l(A)$. We say that X_i are non-degenerate if they are not degenerate. Before we state the main result of the paper we state two simpler results which can be deduced from our method. The following version of the central limit theorem follows from our work:

Theorem 1. *Let X_i, Y_i and x_0 be as above. Suppose that Y_1 has finite second moments and the law of X_i is non-degenerate. Then there is a vector $v_0 \in \mathbf{R}^d$ such that the distribution of $(Y_l - lv_0)/\sqrt{l}$ weakly converges to a Gaussian distribution.*

The limit distribution of course depends on the distributions of X_i . We do not describe this dependence explicitly, but the mean and covariance matrix could be computed from the proof. Here we only mention, that the covariance matrix is invariant under the rotation parts of the elements in the support of X_i .

This result is not new, it is covered by some of the references mentioned earlier. Moreover, our proof is based on Tutubalin's paper. We give this argument in the greater generality discussed in this paper. (Tutubalin assumes that $\text{supp}(X_i)$ generates a dense subgroup of $\text{Isom}(\mathbf{R}^d)$. Moreover, he discusses the cases $d = 2, 3$ only, although he does not seem to use this restriction in an essential way.) In addition, we obtain quantitative bounds which will be necessary for proving the explicit error estimates in Theorem 3 below.

Theorem 2. *Let X_i, Y_i and x_0 be as above. Suppose that Y_0 has finite moments of order $d^2 + 3d + 1$ and X_i are non-degenerate. Let f be any continuous and compactly supported function. Then there is $v_0 \in \mathbf{R}^d$ and $c \in \mathbf{R}$ depending only on the distribution of X_i , such that*

$$\lim_{l \rightarrow \infty} l^{d/2} \mathbf{E}[f(Y_l - lv_0)] = c \int f(x) dx.$$

First we remark that v_0 is the same as in the previous theorem and c can be computed from the covariance matrix of the limit distribution. Moreover, it turns out from the proof, that v_0 is almost surely fixed by the rotation part of X_1 , hence it is 0, if the rotation part of the support of X_1 is sufficiently rich.

When $v_0 = 0$, the Local Limit Theorem can be interpreted as follows: The probability that Y_l belongs to a fixed compact set with smooth boundary is asymptotic to $cl^{-d/2}$ times the Lebesgue measure of the set.

In the local limit theorem, we need the finiteness of high order moments for technical reasons. However, if the group generated by the support of X_i is dense in $\text{Isom}(\mathbf{R}^d)$ then our arguments imply the local limit theorem under the assumption of finite second moments only. In fact, this is true under much weaker assumptions on the group generated by $\text{supp}(X_i)$ (see Theorem 3 below).

Now we formulate the main result of the paper which gives a quantitative description of the distribution of Y_l on multiple scales. However, we need a more restrictive assumption that we call (SSR) . We postpone the definition to the next section, where we explain the notation used through the paper. For now, we only mention that (SSR) holds for example if $\text{supp}(X_i)$ generates a dense subgroup of $\text{Isom}(\mathbf{R}^d)$ and $d \geq 3$. In addition, we can improve the error terms under stronger conditions, i.e. symmetry or (E) . These will be defined in the next section, as well.

Theorem 3. *Let X_i, Y_i and x_0 be as above. Suppose that the law of X_i is non-degenerate, satisfies (SSR) and Y_l has finite moments of order α for some $\alpha > 2$. Furthermore, let f be a smooth function of compact support. Then there is a point $y_0 \in \mathbf{R}^d$, a quadratic form $\Delta(x, x)$ and constants C_Δ and $c > 0$ that depend only on the law of X_i , such that*

$$\begin{aligned} \mathbf{E}[f(Y_l)] &= C_\Delta l^{-d/2} \int f(x) e^{-\Delta(x-y_0, x-y_0)/l} dx \\ &+ O(l^{-\frac{d+\min\{1, \alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}}) \|f\|_1 + O(e^{-cl^{1/4}}) \|f\|_{W^{2, (d+1)/2}}. \end{aligned} \quad (1)$$

In addition, if μ is symmetric or satisfies (E) , we have

$$\begin{aligned} \mathbf{E}[f(Y_l)] &= C_\Delta l^{-d/2} \int f(x) e^{-\Delta(x-y_0, x-y_0)/l} dx \\ &+ O(l^{-\frac{d+\min\{2, \alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}}) \|f\|_1 + O(e^{-cl^{1/4}}) \|f\|_{W^{2, (d+1)/2}}. \end{aligned}$$

The implied constants depend only on the law of X_i .

A few remarks are in order about the conclusion of this theorem. The norm $\|\cdot\|_1$ is the L^1 -norm and $\|\cdot\|_{W^{2, (d+1)/2}}$ is an L^2 Sobolev norm defined by

$$\|f\|_{W^{2, (d+1)/2}}^2 = \int |\widehat{f}(\xi)|^2 (1 + |\xi|)^{d+1} d\xi.$$

The first term on the right hand sides is the main term, it is the integral of f with respect to a Gaussian measure centered at y_0 , and with covariance

matrix Δ/l ; C_Δ is simply a normalizing factor. It will follow from the proof that Δ is invariant under the rotation parts of elements of the support of μ .

The other two terms are error terms. The first is responsible for the large scale, and the second is for the small scale behavior of the random walk. To illustrate this, fix a smooth compactly supported function F , and consider the family $f_l(x) = r_l^{-d}F(x/r_l)$ associated to a sequence of scales r_l . (I.e. the diameter of the support of f_l is proportional to r_l .) It is easily seen that as long as $r_l < \sqrt{l}$, the order of magnitude of the main term is $l^{-d/2}$, the first error term is $l^{-\frac{d+\min\{1,\alpha-2\}}{2}}$ while the second error term is $e^{-cl^{1/4}}r_l^{-d-1/2}$. This shows that the theorem gives a good approximation in the scale range $\sqrt{l} \geq r_l \geq e^{-c'l^{1/4}}$.

The factor $O(e^{-cl^{1/4}})$ is probably not optimal. In fact, our proofs lead to better estimates in some cases. This is discussed in detail in Section 3 after Theorem A, after the necessary background is explained.

Acknowledgments. I am greatly indebted to Elon Lindenstrauss for telling me about the problem and his continued interest in my project. I am also grateful to Jean Bourgain and Emmanuel Breuillard for helpful conversations about various aspects of the problem. I thank Noam Berger, Alex Lubotzky and Nikolay Nikolov for suggesting references [8], [4] and [12] respectively.

I am grateful for the hospitality of the Mathematical Sciences Research Institute, Berkeley, USA.

2 Notation and outline

We identify the isometry group of the d -dimensional Euclidean space with the semidirect product $\text{Isom}(\mathbf{R}^d) = \mathbf{R}^d \rtimes \text{O}(d)$. For $\gamma = (v, \theta) \in \mathbf{R}^d \rtimes \text{O}(d)$ and a point $x \in \mathbf{R}^d$ we write

$$\gamma(x) = v + \theta x,$$

and we define the product of two isometries by

$$(v_1, \theta_1)(v_2, \theta_2) = (v_1 + \theta_1 v_2, \theta_1 \theta_2).$$

If γ is an isometry, we write $v(\gamma)$ for the translation component and $\theta(\gamma)$ for the rotation component of γ in the above semidirect decomposition.

Let μ be a probability measure on $\text{Isom}(\mathbf{R}^d)$. Define the convolution $\mu * \mu$ in the usual way by

$$\int_{\text{Isom}(\mathbf{R}^d)} f(\gamma) d\mu * \mu(\gamma) = \int_{\text{Isom}(\mathbf{R}^d)} \int_{\text{Isom}(\mathbf{R}^d)} f(\gamma_1 \gamma_2) d\mu(\gamma_1) d\mu(\gamma_2),$$

for $f \in C(\text{Isom}(\mathbf{R}^d))$ and write

$$\mu^{*(l)} = \underbrace{\mu * \cdots * \mu}_{l\text{-fold}}$$

for the l -fold convolution. With this notation, $\mu^{*(l)}$ is the distribution of the product of l independent random element of $\text{Isom}(\mathbf{R}^d)$ of law μ . We define the measure $\tilde{\mu}$ by the formula

$$\int_{\text{Isom}(\mathbf{R}^d)} f(\gamma) d\tilde{\mu}(\gamma) = \int_{\text{Isom}(\mathbf{R}^d)} f(\gamma^{-1}) d\mu(\gamma),$$

for $f \in C(\text{Isom}(\mathbf{R}^d))$ and say that μ is symmetric if $\tilde{\mu} = \mu$. The measure μ also acts on measures on \mathbf{R}^d in the following way: If ν is a measure on \mathbf{R}^d , we can define another measure $\mu.\nu$ on \mathbf{R}^d by:

$$\int_{\mathbf{R}^d} f(x) d\mu.\nu(x) = \int_{\text{Isom}(\mathbf{R}^d)} \int_{\mathbf{R}^d} f(\gamma(x)) d\mu(\gamma) d\nu(x),$$

for $f \in C(\mathbf{R}^d)$.

We write δ_{x_0} for the Dirac delta measure concentrated at the point x_0 . With this notation, the law of Y_l , the l th step of the random walk is

$$\mu^{*(l)}.\delta_{x_0}.$$

Write $\theta(\mu)$ for the projection of μ on $O(d)$, i.e. for $f \in C(O(d))$

$$\int_{O(d)} f(\sigma) d\theta(\mu)(\sigma) = \int_{\text{Isom}(\mathbf{R}^d)} f(\theta(\gamma)) d\mu(\gamma).$$

Denote by $G \subset \text{Isom}(\mathbf{R}^d)$ the closure of the group generated by $\text{supp}(\tilde{\mu} * \mu)$. Fix any element $\gamma_0 \in \text{supp} \mu$. Then it is clear that $\text{supp} \mu \subset \gamma_0 G$.

We can replace μ by $\mu' = \mu^{*(k)}$ for some integer $k > 1$ without loss of generality, since Y_{lk+j} , the $lk+j$ th step of the original random walk is the

l th step of the modified random walk started from the random point Y_j . If we do so then we replace G by G' , the closure of the group generated by $\text{supp}(\widetilde{\mu^{*(k)}} * \mu^{*(k)})$. It can be seen easily that G' is the closure of the group generated by

$$G \cup \gamma_0^{-1}G\gamma_0 \cup \dots \cup \gamma_0^{-k+1}G\gamma_0^{k-1}.$$

In Lemma 21 we will see that if we choose k sufficiently large than $\theta(G')$ is normalized by $\theta(\gamma_0^k)$.

Denote by $K \subset O(d)$ the closure group generated $\text{supp} \theta(\tilde{\mu} * \mu)$. By the previous paragraph, we can (and will throughout the paper) assume without loss of generality that K is normalized by $\theta_0 := \theta(\gamma_0)$. Denote by K° the connected component of K .

Denote by μ_K the Haar measure on the group K .

Now we list the various conditions that we will stipulate on μ in various parts of the paper. Some of these were already mentioned in Theorem 3.

(C) ("Centered") The barycenter of the image of the origin in \mathbf{R}^d under μ is the origin, i.e.

$$\int \gamma(0) d\mu(\gamma) = 0.$$

(E) ("Even") The action of K on \mathbf{R}^d is "even", i.e. for every $v \in \mathbf{R}^d$, there is $\theta_v \in K$ such that $\theta_v v = -v$.

(SSR) ("Semi-simple rotations") K° is semi-simple, and there is no non-zero point in \mathbf{R}^d which is fixed by K° .

We also recall the conditions we already defined for convenient reference. We say that μ is *non-degenerate*, if there is no proper closed subset $A \subset \mathbf{R}^d$ and an isometry $\gamma \in \text{Isom}(\mathbf{R}^d)$ such that $\mu^{*(l)}.\delta_{x_0}$ is almost surely contained in $\gamma^l(A)$.

It will be useful for us in many places of the paper to symmetrize μ by replacing it with $\tilde{\mu} * \mu$. Unfortunately, the measure we obtain this way might be degenerate. (For example, consider \mathbf{R}^2 and a measure μ such that $\theta(\text{supp } \mu)$ consists of a single rotation of infinite degree given by a matrix with rational entries.) For this reason, we introduce a different notion which is easily seen to descend to $\tilde{\mu} * \mu$. We say that μ is *almost non-degenerate* if for every point $x \in \mathbf{R}^d$, the set $\{\gamma(x) : \gamma \in \text{supp } \mu\}$ does not lie in a proper affine subspace. As we will see in Lemma 22, if μ is non-degenerate then $\mu^{*(k)}$ is almost non-degenerate for some integer $k \geq 1$. The implication in

the other direction is often true, as well. In particular, almost non-degeneracy is sufficient for most of the paper, except for Section 7.2, where we need to refine this notion.

We say that μ have *finite moments of order $\alpha > 0$* , if

$$\int |v(\gamma)|^\alpha d\mu(\gamma) < \infty.$$

Finally, we say that μ is symmetric if $\mu = \tilde{\mu}$.

A few remarks are in order regarding the role of these conditions. *Non-degeneracy* is clearly necessary for the local limit theorem. However, we can not impose it always for reasons discussed above. On the other hand, almost non-degeneracy is required throughout the paper. Condition *(SSR)* is needed to control the behavior of $\mu^{*(l)}$ on very small scales (up to $e^{-cl^{1/4}}$). Under this assumption we can utilize some powerful results about random walks on semi-simple compact Lie groups. We assume *(SSR)* throughout Section 3 and some other parts of the paper. *Symmetry* or *(E)* allows us to improve the error terms in Theorem 3. They will be assumed in certain parts of Section 4 to show that the cubic terms in certain Taylor expansions cancel with each other. We assume throughout the paper that μ has *finite moments of order 2*. In Section 4 we assume finite moments of order α for $2 \leq \alpha \leq 4$ and the quality of our error terms depend on α . To be able to conclude the local limit theorem in the absence of *(SSR)* we assume the finiteness of higher order moments in Section 7.2. Finally, *(C)* is an assumption which does not restrict generality as we will see in Lemma 20. Therefore we assume it throughout the paper to simplify our arguments.

Now we introduce some further notation and indicate the general strategy of the proof of Theorem 3. Recall that the distribution of the random walk started at the point x_0 after l -steps is the measure

$$\mu^{*(l)}. \delta_{x_0}.$$

It is a simple calculation to check that

$$\mu^{*(l+1)}. \delta_{x_0} = \mu. (\mu^{*(l)}. \delta_{x_0}).$$

Hence our main goal is to understand the operation $\nu \mapsto \mu. \nu$.

This is achieved by studying the Fourier transform, which is given by the formula

$$\widehat{\nu}(\xi) = \int e(\langle \xi, x \rangle) d\nu(x),$$

where $e(x) := e^{-2\pi i x}$. For the Fourier transform of $\mu.\nu$ we get

$$\begin{aligned} (\mu.\nu)^\wedge(\xi) &= \int e(\langle \xi, \gamma(x) \rangle) d\mu(\gamma) d\nu(x) \\ &= \int e(\langle \xi, v(\gamma) + \theta(\gamma)(x) \rangle) d\mu(\gamma) d\nu(x) \\ &= \int e(\langle \xi, v(\gamma) \rangle) \widehat{\nu}(\theta(\gamma)^{-1}\xi) d\mu(\gamma). \end{aligned} \quad (2)$$

This formula shows that the action of μ on the Fourier transform of ν can be disintegrated with respect to spheres centered at the origin.

For every $r \geq 0$, we define a unitary representation of the group $\text{Isom}(\mathbf{R}^d)$ on the space $L^2(S^{d-1})$. Let

$$\rho_r(\gamma)\varphi(\xi) = e(r\langle \xi, v(\gamma) \rangle)\varphi(\theta(\gamma)^{-1}\xi) \quad (3)$$

for $\gamma \in \text{Isom}(\mathbf{R}^d)$, $\varphi \in L^2(S^{d-1})$ and $\xi \in S^{d-1}$.

We also define the operator

$$S_r(\varphi) = \int \rho_r(\gamma)(\varphi) d\mu(\gamma). \quad (4)$$

For a function $\varphi \in C(\mathbf{R}^d)$ and $r \geq 0$, we denote by $\text{Res}_r\varphi$ its restriction to the sphere of radius r . I.e. $\text{Res}_r : C(\mathbf{R}^d) \rightarrow C(S^{d-1})$ is an operator defined by $[\text{Res}_r\varphi](\xi) = \varphi(r\xi)$ for $|\xi| = 1$.

With this notation, we can write (2) as

$$\text{Res}_r(\widehat{\mu.\nu})(\xi) = S_r(\text{Res}_r\widehat{\nu})(\xi).$$

Operators similar to S_r were introduced by Kařdan [16] and Guivarc'h [13]. Guivarc'h proved in the $d = 2$ case, when K is Abelian, that $\|S_r\| < 1 - cr^2$ for $r < 1$ and $\|S_r\| < 1 - c_r$ for $r \geq 1$, where $c > 0$ is a constant depending only on μ , while c_r also depend on r . These estimates are sufficient for proving a ratio limit theorem, and as Breuillard [5] pointed out, combined with the central limit theorem, it is sufficient even for a local limit theorem. We are unable to prove such strong estimates, but we will prove in Section 3 a weaker version: Proposition 4 which is still sufficient for our application. In brief, we prove the estimate with constants c and c_r which (mildly) depend on the oscillations of φ . The proof is based on mixing properties of random walks on semi-simple compact Lie groups (see Theorem A below).

Using the estimates given in Section 3, we can show that the Fourier transform of the random walk after l steps "lives in" the ball of radius $l^{-1/2} \log l$. These estimates alone are sufficient for the ratio limit theorem but not for the central or local limit theorems. The frequency range $r > l^{-1/2} \log l$ is responsible for the second error term in Theorem 3.

We need a more precise understanding of the Fourier transform of $\mu^{*(l)}.\delta_{x_0}$ in the range $r \leq l^{-1/2} \log l$. This frequency range contributes the main term and the first error term in Theorem 3. In section 4 we give Tutubalin's [23] argument for the central limit theorem in the more general setting that we consider and obtain the explicit estimates. In brief, this argument is based on decomposing $L^2(S^{d-1})$ as the orthogonal sum of several subspaces and using the Taylor expansion of the function $e(x)$ showing that these subspaces are almost invariant for S_r . We show that the contribution of only one of these subspaces is significant and that on this subspace rotations act trivially. Hence the problem is reduced to the easy case of sum of independent random variables.

We will encounter L^2 spaces on various submanifolds of \mathbf{R}^d . We always consider them with respect to the "natural" measure, i.e. which is invariant under isometries. When the manifold is compact we normalize the measure to be probability.

Throughout the paper the letters c, C and various subscripted versions refer to constants and parameters. The same symbol occurring in different places need not have the same value unless the contrary is explicitly stated. For convenience, we use lower case for constants which are best thought of to be small and upper case for those which are best thought of to be large. In addition, we occasionally use Landau's O and o notation.

The organization of the rest of the paper that we have not explained yet is as follows: In Section 5, we combine the estimates of Section 4 and 3 to conclude Theorem 3. In Section 6 we derive Theorem 1 as a corollary of the results in Section 4. Finally, in Section 7 we prove Theorem 2. When K is Abelian and its action on \mathbf{R}^d has a trivial component some additional difficulties arise which prevents us from using the method of Guivarc'h [13]. To address these issues in Section 7.2, we use Taylor expansions motivated by Tutubalin's paper. For this argument we need to assume the finiteness of high order moments.

3 Estimates for high frequencies

The goal of this section is to estimate the norm of the operator S_r defined in Section 2. We are not able to show that $\|S_r\| < 1$, but we can give the following estimate for $\|S_r\varphi\|_2$ in terms of the Lipschitz norm of φ .

Proposition 4. *Suppose that μ is almost non-degenerate, has finite moments of order 2, and satisfies (SSR) and (C). Then there is a constant $c > 0$ depending only on μ such that the following hold. Let $\varphi \in \text{Lip}(S^{d-1})$ with $\|\varphi\|_2 = 1$. Then*

$$\|S_r\varphi\|_2 \leq 1 - c \min\left\{r^2, \frac{1}{\log^3(\|\varphi\|_{\text{Lip}} + 2)}\right\}. \quad (5)$$

This estimate allows us to control the Fourier transform of the random walk in the frequency range $e^{cl^{1/4}} > r > l^{-1/2} \log l$.

We warn the reader that the proof of the proposition depends on the central limit theorem, (Theorem 1); this is in fact the only place where almost non-degeneracy is used. We point out that the central limit theorem is proved using only Proposition 10 which is independent of the arguments of the current section. In particular, Proposition 19 which is a refinement of Proposition 10 based on the results of this section is not needed.

Mixing properties of random walks on semi-simple compact Lie groups is a crucial ingredient of our proof. We state the result that we use in the next theorem. The proof will be given in the forthcoming paper [24, Corollary 7]. A quantitatively weaker version, but essentially sufficient for our purpose could be deduced from the Solovay-Kitaev algorithm, at least in the case $K = \text{SU}(d)$. The Solovay-Kitaev algorithm was first described in an e-mail discussion list by Solovay in 1995. Kitaev independently discovered it and published it in 1997 [19]. For a recent exposition see [10]. See also the paper of Dolgopyat [11, Theorems A.2 and A.3], which provides similar estimates.

Theorem A. *Let K be a compact Lie group with semi-simple connected component. Let μ be a symmetric probability measure on K such that $\text{supp}(\mu)$ generates a dense subgroup in K . Then there is a constant $c > 0$ depending only on μ such that the following hold. Let $\varphi \in \text{Lip}(K)$ be a function such that $\|\varphi\|_2 = 1$ and $\int \varphi d\mu_K = 0$. Then*

$$\left\| \int \varphi(\theta^{-1}\sigma) d\mu(\theta) \right\|_2 < 1 - \frac{c}{\log^2(\|\varphi\|_{\text{Lip}} + 2)} \quad (6)$$

Recall that m_K denotes the Haar measure on K .

We mention that Bourgain and Gamburd [6], [7] proved (in the $K = \mathrm{SU}(d)$ case) that if μ satisfies some additional conditions (i.e. the support of μ consists of matrices with algebraic entries) then the right hand side is estimated by $1 - c$ independently of φ . As we mentioned in the introduction, Conze and Guivarc'h [9] proved the ratio limit theorem under certain assumption. This assumption is that $K = \mathrm{SO}(d)$, and $\theta(\mu)$ satisfies (6) with $1 - c$ on the right independently of φ .

If one improves the estimate in Theorem A, then our argument presented below provides better estimates in Proposition 4 and Theorem 3. In particular, one can replace the right hand side of (6) with $1 - c \log^{-A}(\|\varphi\|_{\mathrm{Lip}} + 2)$, then one can write $1 - c \min\{r^2, \log^{-A-1}(\|\varphi\|_{\mathrm{Lip}} + 2)\}$ on the right hand side of (5) and $O(e^{-cl^{1/(A+2)}})\|f\|_W$ instead of the second error term in (1). In fact, Theorem A is proved with better bounds for most Lie groups; except for those which project onto $\mathrm{SO}(3)$. For details, we refer to [24]. Moreover, for certain generators (e.g. when they are given with algebraic entries), the estimates are available even with $A = 0$, (as we discussed above).

The rest of the section is devoted to the proof of Proposition 4. A simple observation shows that it is enough to prove it for symmetric measures. Indeed, we have

$$\|S_r \varphi\|_2^2 = \langle S_r \varphi, S_r \varphi \rangle = \langle \varphi, S_r^* S_r \varphi \rangle \leq \|S_r^* S_r \varphi\|_2 \quad (7)$$

and $S_r^* S_r$ is the operator analogous to S_r corresponding to the symmetric measure $\tilde{\mu} * \mu$. It is easy to check that if any of almost non-degeneracy, finite moments of order α or (SSR) holds for μ , then the same conditions also holds for $\tilde{\mu} * \mu$. It may happen that (C) fails for $\tilde{\mu} * \mu$, but we can change the origin so that it will hold. (See Lemma 20.) In addition, this argument shows that we can assume that S_r is positive. From now on, until the end of the section, we assume, that μ is symmetric, almost non-degenerate, has finite second moments and satisfies (C) and (SSR). Moreover we assume that S_r is selfadjoint and positive.

We begin with a simple lemma which shows that the length of the translation part of γ with $\mu^{*(l)}$ probability at least $9/10$ is bounded above by $\ll \sqrt{l}$.

Lemma 5.

$$\int |v(\gamma)|^2 d\mu^{*(l)}(\gamma) = l \cdot \int |v(\gamma)|^2 d\mu(\gamma).$$

Proof. For $l = 1$ the statement is obvious, for l larger the proof is by induction:

$$\begin{aligned}
\int |v(\gamma)|^2 d\mu^{*(l+1)}(\gamma) &= \int \int |v(\gamma_1 \gamma_2)|^2 d\mu(\gamma_2) d\mu^{*(l)}(\gamma_1) \\
&= \int \int |v(\gamma_1) + \theta(\gamma_1)v(\gamma_2)|^2 d\mu(\gamma_2) d\mu^{*(l)}(\gamma_1) \\
&= \int \int |\theta(\gamma_1^{-1})v(\gamma_1) + v(\gamma_2)|^2 d\mu(\gamma_2) d\mu^{*(l)}(\gamma_1) \\
&= \int \int (|v(\gamma_1)|^2 + |v(\gamma_2)|^2 \\
&\quad + 2\langle \theta(\gamma_1^{-1})v(\gamma_1), v(\gamma_2) \rangle) d\mu(\gamma_2) d\mu^{*(l)}(\gamma_1).
\end{aligned}$$

Integrating out γ_2 , the third term in the last line vanishes by (C). This proves the lemma by induction. \square

Until the end of the section, we fix $r > 0$ and a function $\varphi \in \text{Lip}(S^{d-1})$ and prove Proposition 4 for these. The strategy of the proof is the following. We fix two integers

$$l_1 = [C_1(r^{-2} + \log^3(\|\varphi\|_{\text{Lip}} + 2))], \quad l_2 = [C_2(r^{-2} + \log^3(\|\varphi\|_{\text{Lip}} + 2))]$$

and where C_1, C_2 are suitably chosen large constants depending on μ but not on φ or r .

Consider the set of isometries, for which φ is an almost invariant vector in the ρ_r representation. More precisely, define the set

$$B(\varepsilon) := \{\gamma \in \text{Isom}(\mathbf{R}^d) : \|\rho_r(\gamma)\varphi - \varphi\|_2 < \varepsilon\}.$$

In what follows, we assume to the contrary that

$$\mu^{*(l_i)}(B(\varepsilon)) > 9/10 \tag{8}$$

for both $i = 1$ and $i = 2$, for some $\varepsilon > 0$ whose value will be determined at the end of the proof depending on μ .

Now we clarify the dependence of the three parameters ε , C_1 and C_2 we introduced. First ε is to be set to be sufficiently small depending on μ , then we set C_1 to be sufficiently large depending on μ and ε , finally we specify the value of C_2 sufficiently large depending on C_1 and μ .

At the end of the section, we will see that the failure of (8) indeed implies Proposition 4, and the constant c in the proposition depends only on ε , C_1 and C_2 .

We proceed by various Lemmata which show under the hypothesis (8) that $B(\varepsilon')$ contains larger and larger families of isometries if ε' becomes larger and larger. We will reach contradiction when we show that we can find translations of length comparable to r^{-1} in many directions.

In the first lemma, we conclude that $B(4\varepsilon)$ contains isometries with an arbitrary prescribed rotation part and translation part proportional to \sqrt{l} .

Lemma 6. *Suppose that (8) holds with some $\varepsilon > 0$ for $i = 1$. Suppose further that C_1 is sufficiently large depending on μ and ε . Then there exist a constant C which depends only on μ and C_1 such that the following holds: There is a set $X \subset B(4\varepsilon)$ such that*

$$\begin{aligned}\theta(X) &= K \\ |v(\gamma)| &< C\sqrt{l_1} \quad \text{for } \gamma \in X.\end{aligned}$$

Proof. We deduce the lemma from Theorem A. Let \mathcal{B} be the $c\|\varphi\|_{\text{Lip}}^{-1}$ neighborhood of the identity in K° . The choice of $c > 0$ depends on ε and the geometry of $K^\circ \subset \text{O}(d)$, and we do it in such a way that any element $\sigma \in \mathcal{B}$ moves a point of S^{d-1} to distance at most $\varepsilon\|\varphi\|_{\text{Lip}}^{-1}$. Then we have $\mathcal{B} \subset B(\varepsilon)$.

Take an approximate identity ψ on K , which has the following properties:

$$\text{supp}(\psi) \subset \mathcal{B}, \quad \int \psi dm_K = 1 \quad \text{and} \quad \|\psi\|_{\text{Lip}} \leq C\|\varphi\|_{\text{Lip}}^{1+\dim K}.$$

Note that these imply that $\|\psi\|_2 \leq C\|\varphi\|_{\text{Lip}}^{(\dim K)/2}$. These constants again depend only on K and ε . Now we apply Theorem A successively l_1 times starting with the function $(1 - \psi)/\|1 - \psi\|_2$, and get

$$\|1 - \int \psi(\theta^{-1}\sigma) d\theta(\mu)^{* (l_1)}(\theta)\|_2 \leq \frac{1}{10}$$

provided C_1 is sufficiently large depending only on μ and ε . Recall that $l_1 > C_1 \log^3(\|\varphi\|_{\text{Lip}} + 2)$. We note that taking the average of translates of ψ may only decrease the Lipschitz norm.

Now let $Y \subset B(\varepsilon)$ be such that $\mu^{*(l_1)}(Y) > 8/10$, and

$$|v(\gamma)| < C\sqrt{l_1}/2 \quad \text{for } \gamma \in Y.$$

For a sufficiently large C depending on (the second moment of) μ and C_1 , this is possible due to the assumption (8) and Lemma 5. Denote by ν the measure we obtain from $\theta(\mu^{*(l_1)})$ if we restrict it to the set Y , and normalize it to get a probability measure. Then we have

$$\left\| \int \psi(\theta^{-1}\sigma) d\nu(\theta) \right\|_2 \leq \left(1 + \frac{1}{10}\right) \cdot \frac{10}{8} < \sqrt{2}.$$

Thus

$$m_K(\theta(Y)\mathcal{B})) \geq m_K(\text{supp}(\int \psi(\theta^{-1}\sigma) d\nu(\theta))) > \frac{1}{2},$$

which proves the lemma with the choice $X = \theta(Y)\mathcal{B}\theta(Y)\mathcal{B}$. \square

Lemma 7. *There are constants $c, C > 0$ which depend only on μ such that, we have*

$$\mu^{*(l)}(\gamma : |\langle v(\gamma), u_0 \rangle| > c\sqrt{l} \text{ and } |v(\gamma)| < C\sqrt{l}) > 1/2,$$

for any sufficiently large (depending only on μ) integer l , and any $u_0 \in S^{d-1}$.

Proof. The lemma is an easy consequence of the central limit theorem, i.e. Theorem 1. \square

We could, off course, replace $1/2$ in the lemma with any number less than 1.

Now we can show that under our standing assumption (8), $B(9\varepsilon_1)$ contains a nontrivial translation.

Lemma 8. *Suppose that (8) holds with some $1/2 > \varepsilon > 0$ for $i = 1$ and $i = 2$. Suppose further that C_1 is sufficiently large so that Lemma 6 holds and C_2 is sufficiently large depending on μ and C_1 . Then there are constants $c, C > 0$ depending only on μ , such that for any $u_0 \in S^{d-1}$, there is an element $\gamma_1 \in B(5\varepsilon)$ with the following properties:*

$$|v(\gamma_1)| < C\sqrt{l_2}, \quad \langle v(\gamma_1), u_0 \rangle > c\sqrt{l_2} \quad \text{and} \quad \theta(\gamma_1) = 1.$$

Proof. Denote by c_0 and C_0 the constants from Lemma 7. Using that lemma with $l = l_2$ and (8) with $i = 2$, we find an element $\gamma_2 \in B(\varepsilon)$ such that

$$|\langle v(\gamma_2), u_0 \rangle| > c_0\sqrt{l_2} \quad \text{and} \quad |v(\gamma_2)| < C_0\sqrt{l_2}.$$

On the other hand, applying to Lemma 6, we can find $\gamma_3 \in B(4\varepsilon)$ that satisfies:

$$\theta(\gamma_3) = \theta(\gamma_2)^{-1} \quad \text{and} \quad |v(\gamma_3)| < C'_0 \sqrt{l_1},$$

where C'_0 is the constant C from that lemma.

We demand that $l_2/l_1 = C_2/C_1$ is so large that

$$c_0 \sqrt{l_2} > 2C'_0 \sqrt{l_1}.$$

Then it is an easy calculation to verify that $\gamma_1 = \gamma_2 \gamma_3$ has the claimed properties. □

Recall the definition of l_2 , in particular that it implies $\sqrt{l_2} > \sqrt{C_2} r^{-1}$. The next Lemma shows that we can find a translation γ'_1 with properties similar to that of γ_1 in the previous lemma, but which is shorter.

Lemma 9. *Under the same hypothesis as in Lemma 8, there is a constant $c > 0$ which depend only on μ , and there is an element $\gamma'_1 \in B(26\varepsilon)$ with the following properties:*

$$|v(\gamma'_1)| < r^{-1}/2, \quad \langle v(\gamma'_1), u_0 \rangle > cr^{-1} \quad \text{and} \quad \theta(\gamma'_1) = 1.$$

Proof. Let $\gamma_1 \in B(5\varepsilon)$ be an isometry with the properties stated in Lemma 8, and write $v = v(\gamma_1)$. For simplicity we assume that $\langle v, u_0 \rangle > 0$; the other case is similar. By the assumption (SSR), we have

$$\int \langle \theta v, u_0 \rangle dm_{K^\circ}(\theta) = 0.$$

Thus, there is $\theta_1 \in K^\circ$, such that $\langle \theta_1 v, u_0 \rangle \leq 0$.

There is a curve $\Theta : [0, 1] \rightarrow K^\circ$ such that $\Theta(0) = 1$ and $\Theta(1) = \theta_1$, and the length of the curve $[0, 1] \rightarrow \Theta(t)v$ is less than $C|v|$, where C depends only on the embedding of K° to $O(d)$, hence on μ . Then there is a sequence of rotations

$$\sigma_0 = 1, \sigma_1, \sigma_2, \dots, \sigma_N = \theta_1 \in K^\circ$$

with $N \leq 2Cr|v| + 1$ such that for any $1 \leq i \leq N$

$$|\sigma_i v - \sigma_{i-1} v| < r^{-1}/2.$$

By the pigeon hole principle, there is an index $1 \leq i \leq N$, such that

$$\langle \sigma_{i-1}v - \sigma_i v, u_0 \rangle \geq \langle v, u_0 \rangle / N \geq cr^{-1}$$

with a suitably small constant $c > 0$. (c depends on C and the constants appearing in the previous lemma.)

Now let $g_i \in B(4\varepsilon)$ be such that $\theta(g_i) = \sigma_i$; such elements can be found by virtue of Lemma 6. The proof is finished by an easy verification of the stated properties for the element

$$\gamma'_1 := g_{i-1}\gamma_1 g_{i-1}^{-1} g_i \gamma_1^{-1} g_i^{-1}.$$

□

Proof of Proposition 4. Fix a function φ and r . We assume to the contrary that (8) holds for $i = 1$ and $i = 2$.

Now let $c > 0$ be the constant from Lemma 9. Clearly, there is a point $u_0 \in S^{d-1}$, such that

$$\int_{|\xi - u_0| < c/2} |\varphi|^2 d\xi > c', \quad (9)$$

with a constant $c' > 0$ that depends only on c and d . By Lemma 9, there is an element $\gamma'_1 \in B(26\varepsilon)$ such that $\theta(\gamma'_1) = 1$ and $v' := v(\gamma'_1)$ satisfies $|v'| < r^{-1}/2$, and $|\langle v', u_0 \rangle| > cr^{-1}$. This leads to the inequality

$$\int_{S^{d-1}} |(1 - e(r\langle \xi, v' \rangle))\varphi(\xi)|^2 d\xi < (26\varepsilon)^2.$$

If $|\xi - u_0| < c/2$, then $c/2 < |r\langle \xi, v' \rangle| < 1/2$. This and (9) gives

$$|1 - e(c/2)|^2 c' < (26\varepsilon)^2,$$

which is a contradiction if we choose ε to be sufficiently small. Since c and c' depends only on μ , it follows that ε depends only on μ . We chose C_1 depending on μ and ε in Lemma 6 and C_2 depending on C_1 and μ in Lemma 8. Thus all this parameters depend only on μ .

Thus far, we showed that (8) can not hold for both $i = 1$ and $i = 2$, i.e.

$$\mu^{*(l_i)}(B(\varepsilon)) < 9/10,$$

for either $i = 1$ or $i = 2$. For simplicity, assume that l_i is even. Then we can write

$$\begin{aligned}
\|S_r^{l_i/2}(\varphi)\|_2^2 &= \langle S_r^{l_i/2}(\varphi), S_r^{l_i/2}(\varphi) \rangle = \langle S_r^{l_i}(\varphi), \varphi \rangle \\
&= \int \langle \rho_r(\gamma)\varphi, \varphi \rangle d\mu^{*(l_i)}(\gamma) \\
&\leq \int \langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi, \varphi \rangle / 2 d\mu^{*(l_i)}(\gamma) \\
&\leq \int_{\text{Isom}(\mathbf{R}^d) \setminus B(\varepsilon)} \langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi, \varphi \rangle / 2 d\mu^{*(l_i)}(\gamma) + \mu^{*(l_i)}(B(\varepsilon)) \\
&\leq (1 - \varepsilon^2/2)/10 + 9/10.
\end{aligned}$$

To deduce the last inequality, we used the identity

$$\langle \rho_r(\gamma)\varphi + \rho_r(\gamma^{-1})\varphi, \varphi \rangle = 2 - \|\rho_r(\gamma)\varphi - \varphi\|_2^2.$$

We concluded that $\|S_r^{l_i/2}(\varphi)\|_2 \leq e^{-c}$ for some $c > 0$ depending on μ . By selfadjointness of S_r we can deduce that $\|S_r(\varphi)\|_2 \leq e^{-2c/l_i}$. We compare this with the definition of

$$l_i = [C_i(r^{-2}, \log^3(\|\varphi\|_{\text{Lip}} + 2))]$$

and finish the proof. \square

4 Estimates for low frequencies

As in the other sections, μ is a fixed probability measure on $\text{Isom}(\mathbf{R}^d)$. Recall that K is the closure of the rotation group generated by $\text{supp } \theta(\tilde{\mu} * \mu)$. Moreover, we assume that $\text{supp } \theta(\mu) \subset \theta_0 K$, where $\theta_0 \in O(d)$ is a rotation which normalizes K .

Fix a point $x_0 \in \mathbf{R}^d$, the starting point of the random walk, and fix a real number $r \geq 0$. Define

$$\psi_0(\xi) = e(r\langle x_0, \xi \rangle) = \text{Res}_r(\widehat{\delta_{x_0}})(\xi)$$

for $\xi \in S^{d-1}$ which is the Fourier transform of the measure δ_{x_0} restricted to the sphere of radius r . Our objective in this section is to estimate $S_r^l \psi_0$. The estimate will be useful in the range $r < l^{-1/2} \log l$, i.e. when the frequency is sufficiently small.

The next proposition shows that $S_r^l \psi_0$ is approximated by the Fourier transform of a Gaussian distribution with covariance matrix Δ which depends on μ .

Proposition 10. *Assume that μ is almost non-degenerate, has finite moments of order α for some $\alpha \geq 2$ and satisfies (C). There is a constant C and a symmetric positive definite quadratic form $\Delta(\xi, \xi)$ on \mathbf{R}^d invariant under the action of K and θ_0 , such that the following holds*

$$\|S_r^l \psi_0 - e^{-r^2 l \Delta}\|_2 < C(r^{\min\{1, \alpha-2\}} + |x_0|^2 r^2).$$

Moreover, if μ is symmetric or satisfies (E), then we have the better bound:

$$\|S_r^l \psi_0 - e^{-r^2 l \Delta}\|_2 < C(r^{\min\{2, \alpha-2\}} + |x_0|^2 r^2).$$

When $\alpha = 2$, the constant C can be taken arbitrarily small as $r \rightarrow 0$.

The rest of this section is devoted to the proof of this proposition, and in Section 4.4 we will give a slight improvement for the range $r > l^{-1/2}$. However, this improvement requires the assumption (SSR) and it is based on the results of Section 3.

Throughout this section (i.e. Section 4) we make the following assumptions. We assume that r is small, i.e. $r < c \min\{1, |x_0|^{-1}\}$, where c is a suitable small constant. For r larger, the statement of the proposition is vacuous. We assume that μ is almost non-degenerate, has finite moments of order $\alpha \geq 2$ and satisfies (C).

In addition, at certain parts we assume that μ is symmetric or satisfies (E), but we always mention these explicitly.

The argument is based on Tutubalin's paper [23]. The most significant difference is that we consider the following, more general, decomposition of the space $\mathcal{H} := L^2(S^{d-1})$. This is due to the fact that we do not assume $K = \text{SO}(d)$.

Let \mathcal{H}_0 be the subspace of functions $\varphi \in \mathcal{H}$ which are fixed by the action of K , i.e. $\varphi(\theta\xi) = \varphi(\xi)$ for every $\theta \in K$. For later reference we note that if $\varphi \in \mathcal{H}$, then the orthogonal projection of φ to \mathcal{H}_0 is obtained by the formula:

$$\int \varphi(\theta\xi) dm_K(\theta). \tag{10}$$

Denote by $\mathcal{P}_k \subset \mathcal{H}$ the space functions which are restrictions of degree k polynomials to S^{d-1} . We define the spaces \mathcal{H}_k , $k \geq 1$ recursively. Once \mathcal{H}_k

is defined, let \mathcal{H}_{k+1} be the orthogonal complement of \mathcal{H}_k in the space

$$\text{span} \{ \psi \varphi : \psi \in \mathcal{P}_{k+1}, \varphi \in \mathcal{H}_0 \},$$

where $\text{span} \{ \cdot \}$ denotes the smallest closed subspace that contains the functions inside the brackets. Since $\mathcal{P}_k \subset \mathcal{H}_0 \oplus \dots \oplus \mathcal{H}_k$, we have indeed

$$\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \dots$$

In the special case $K = \text{SO}(d)$, \mathcal{H}_k is the familiar space of spherical harmonics of degree k , which was considered in Tutubalin's paper.

Denote by $\mathcal{H}_\infty = \mathcal{H}_4 \oplus \mathcal{H}_5 \oplus \dots$. Finally, let $P_i : \mathcal{H} \rightarrow \mathcal{H}_i$ be the orthogonal projection operator for each $i \in \{0, 1, \dots, \infty\}$.

Taking $r = 0$, it is easy to see that the above subspaces are invariant for S_0 . Below, we will show that they are "almost invariant" for small r ; more precisely, we will bound the norm of $P_i S_r P_j$ by a polynomial of r , for $i \neq j$. Additionally, we will see that the norm of $P_i S_r P_i$ for all $1 \leq i < \infty$ is strictly less than 1. However, the dependence on i would require a more careful analysis. Fortunately, we do not need to do this here, since as we will see, the contribution of the spaces \mathcal{H}_i , $i \geq 4$ is negligible compared to other error terms, and this is why we introduced the space \mathcal{H}_∞ . These estimates, which are simply based on Taylor expansion, will be given in Section 4.1.

To simplify notation, we write $\psi_l = (P_0 S_r P_0)^l \psi_0$ for $l \geq 1$. We will use the almost invariance of S_r mentioned in the previous paragraph to show that ψ_l is a good approximation to $S_r^l \psi_0$. This is done in two steps. We set $P = P_0 + P_1 + P_2 + P_3$, and consider another sequence, defined by $\psi'_l = (P S_r P)^l \psi_0$ for $l \geq 0$. The next two Lemmata that will be proved in Section 4.2, claims that π'_l approximates $S_r^l \psi_0$ and ψ_l approximates ψ'_l .

Lemma 11. *There is a constant $C > 0$ such that the following holds:*

$$\|\psi'_l - S_r^l \psi_0\|_2 \leq C(r^{\min\{\alpha-2, 2\}} + (|x_0|r)^4).$$

Moreover, when $\alpha = 2$, we can replace C in the above estimate by $C_r \rightarrow 0$ as $r \rightarrow 0$.

Lemma 12. *There are constants $C, c > 0$ depending only on μ such that the following holds for $l \geq C \log(r^{-1}|x_0| + 2)$:*

$$\|\psi_l - \psi'_l\|_2 \leq C e^{-cr^2 l} r.$$

If μ is symmetric, then

$$\|\psi_l - \psi'_l\|_2 \leq Ce^{-cr^2l}r^2.$$

If μ satisfies (E), (but not necessarily symmetric), then

$$\|\psi_l - \psi'_l\|_2 \leq Ce^{-cr^2l}(r^{\min\{\alpha-1,2\}} + |x_0|r^2).$$

In light of these Lemmata, it remains to understand the operator $P_0S_rP_0$. This is essentially a multiplication operator as the next formula shows. For $\varphi \in \mathcal{H}_0$:

$$\begin{aligned} P_0S_rP_0\varphi(\xi) &= \int \int e(r\langle\sigma\xi, v(\gamma)\rangle)\varphi(\theta(\gamma)^{-1}\sigma\xi)d\mu(\gamma)dm_K(\sigma) \\ &= F(\xi)\varphi(\theta_0^{-1}\xi), \end{aligned} \tag{11}$$

where

$$F(\xi) = \int \int e(r\langle\sigma\xi, v(\gamma)\rangle)d\mu(\gamma)dm_K(\sigma). \tag{12}$$

Recall that $\text{supp } \theta(\mu) \subset \theta_0K$, and θ_0 normalizes K .

Based on this formula and the Taylor expansion of the function F , we will show the following lemma in Section 4.3.

Lemma 13. *There are constants $C, c > 0$ and a quadratic form Δ on \mathbf{R}^d depending only on μ such that*

$$\|\psi_l - e^{-lr^2\Delta}\|_2 < Ce^{-clr^2}(r^{\min\{1,\alpha-2\}} + |x_0|^2r^2).$$

Δ is invariant under K and θ_0 .

Moreover, if μ is symmetric or satisfies (E), then we have the better estimate

$$\|\psi_l - e^{-lr^2\Delta}\|_2 < Ce^{-clr^2}(r^{\min\{2,\alpha-2\}} + |x_0|^2r^2).$$

Proposition 10 immediately follows from Lemmata 11–13.

4.1 Taylor expansion, and approximate invariance

In this section we give some estimates for the norm of the operators $P_iS_rP_j$. These will be deduced from the following lemma, which is based on Taylor series expansion of the function $e(r\langle\xi, v(\gamma)\rangle)$ which is the multiplier in the ρ_r representation.

Lemma 14. *There is an absolute constant $C > 0$ such that for any $\varphi \in \mathcal{H}$ with $\|\varphi\|_2 = 1$ and $\gamma \in \text{Isom}(\mathbf{R}^d)$ with $\theta(\gamma) \in \theta_0 K$ we have*

$$\|P_i \rho_r(\gamma) P_i \varphi - \rho_0(\gamma) P_i \varphi\|_2 < Cr|v(\gamma)|, \quad (13)$$

$$\|P_j \rho_r(\gamma) P_i \varphi\|_2 < \min\{1, C(r|v(\gamma)|)^{|i-j|}\}. \quad (14)$$

Proof. For the proof, we can assume that $\varphi \in \mathcal{H}_i$. By Taylor's theorem,

$$\begin{aligned} \rho_r(\gamma) \varphi(\xi) &= e(r\langle \xi, v(\gamma) \rangle) \varphi(\theta(\gamma)^{-1} \xi) \\ &= \left[\sum_{m=0}^{M-1} C_m \langle \xi, v(\gamma) \rangle^m + O(r^M |v(\gamma)|^M) \right] \varphi(\theta(\gamma)^{-1} \xi), \end{aligned}$$

where $C_0 = 1$, and C_m are constants depending on m and r , while the implied constant is absolute.

To deduce (13), take $M = 1$, and note that

$$\rho_0(\gamma) \varphi(\xi) = \varphi(\theta(\gamma)^{-1} \xi) = P_i(\varphi(\theta(\gamma)^{-1} \xi)).$$

Since $\varphi \in \mathcal{H}_i$, we have

$$\varphi = p_1 \psi_1 + \dots + p_k \psi_k$$

with some $p_1, \dots, p_k \in \mathcal{P}_i$ and $\psi_1, \dots, \psi_k \in \mathcal{H}_0$. To deduce (14) when $j > i$, take $M = j - i$. Write

$$q(\xi) = \sum_{m=0}^{j-i-1} C_m \langle \xi, v(\gamma) \rangle^m \in \mathcal{P}_{j-i-1}.$$

Then

$$\sum_{m=0}^{j-i-1} C_m \langle \xi, v(\gamma) \rangle^m \varphi(\theta(\gamma)^{-1} \xi) = q(\xi) [p_1(\theta(\gamma)^{-1} \xi) \psi_1(\theta_0^{-1} \xi) + \dots + p_k(\theta(\gamma)^{-1} \xi) \psi_k(\theta_0^{-1} \xi)],$$

where $q(\xi) p_n(\theta(\gamma)^{-1} \xi) \in \mathcal{P}_{j-1}$ for $1 \leq n \leq k$. Thus after applying P_j , this term vanishes. This proves (14) when $j > i$.

Let now $j < i$, and let $\psi \in \mathcal{H}_j$ with $\|\psi\|_2 = 1$ be such that

$$\|P_j \rho_r(\gamma) \varphi\|_2 = \langle \rho_r(\gamma) \varphi, \psi \rangle = \langle \varphi, \rho_r(\gamma^{-1}) \psi \rangle \leq \|P_i \rho_r(\gamma^{-1}) \psi\|_2.$$

Then the claim follows from (14) applied for ψ and γ^{-1} and the role of i and j reversed. \square

Lemma 15. *There are constants $c < 1$ and C depending only on μ such that the following hold*

$$\begin{aligned}\|P_i S_r P_i\| &\leq c \quad \text{for } r < c \quad \text{and } 1 \leq i \leq 3, \\ \|P_i S_r P_j\| &\leq C r^{\min\{|i-j|, \alpha\}}.\end{aligned}$$

Moreover, when $|i-j| > \alpha$, we can replace the second estimate by $o(r^{\min\{|i-j|, \alpha\}})$.

The significance of the last line is simply that we can conclude the central limit theorem even in the case $\alpha = 2$.

Proof. To prove the first inequality, we integrate (13) with respect to $d\mu(\gamma)$:

$$\begin{aligned}\|P_i S_r P_i \varphi - S_0 P_i \varphi\|_2 &= \left\| \int P_i \rho_r(\gamma) P_i \varphi - \rho_0(\gamma) P_i \varphi d\mu(\gamma) \right\|_2 \\ &< C r \int |v(\gamma)| d\mu(\gamma).\end{aligned}$$

This estimate shows that it is enough to estimate the norm of S_0 on $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$. Denote by \mathcal{P}' the orthogonal complement of $\mathcal{H}_0 \cap \mathcal{P}_3$ in \mathcal{P}_3 . For each $\varphi \in \mathcal{P}'$, $\|S_0^* S_0 \varphi\|_2 < \|\varphi\|_2$, because otherwise φ would be invariant under K , i.e. $\varphi \in \mathcal{H}_0$. Since \mathcal{P}' is finite dimensional, there is a constant $c < 1$ such that $\|S_0^* S_0 \varphi\|_2 < c \|\varphi\|_2$ for $\varphi \in \mathcal{P}'$ (cf. (7)).

Note that any $\varphi \in \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$ is of the form $p_1 \psi_1 + \dots + p_k \psi_k$ with $p_i \in \mathcal{P}'$ and $\psi_i \in \mathcal{H}_0$. Let $\varphi_1, \dots, \varphi_k$ be an orthogonal basis of \mathcal{P}' consisting of eigenvalues of $S_0^* S_0$. Observe that the spaces $\varphi_i \cdot \mathcal{H}_0$ are eigenspaces of $S_0^* S_0$ with the same eigenvalues as φ_i and they span $\mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3$. Hence we are able to conclude that

$$\|P_i S_0 P_i\|^2 \leq \|P_i S_0^* S_0 P_i\| < c$$

for $i = 1, 2, 3$. This combined with the first formula of the proof gives the first claim provided r is sufficiently small.

For the second claim, we integrate (14):

$$\|P_j S_r(\gamma) P_i\| < \int \min\{1, C(r|v(\gamma)|)^{|i-j|}\} d\mu(\gamma).$$

If $|i-j| \leq \alpha$, we can simply write

$$\|P_j S_r(\gamma) P_i\| < C r^{|i-j|} \int |v(\gamma)|^{|i-j|} d\mu(\gamma),$$

and the claim follows from the moment condition on μ . If $|i - j| > \alpha$, we write

$$\|P_j S_r(\gamma) P_i\| < C r^\alpha \int |v(\gamma)|^\alpha \min\{(r|v(\gamma)|)^{-\alpha}, (r|v(\gamma)|)^{|i-j|-\alpha}\} d\mu(\gamma).$$

Observe that

$$\min\{(r|v(\gamma)|)^{-\alpha}, (r|v(\gamma)|)^{|i-j|-\alpha}\} \leq 1$$

and that it tends to 0 for all γ as $r \rightarrow 0$. Now the claim follows by the dominated convergence theorem. \square

The bound on $P_1 S_r P_0$ in Lemma 15 is not optimal. Indeed, it is easy to see that the linear terms in the Taylor expansions cancel due to condition (C).

Lemma 16. *There is a constant C such that*

$$\|P_1 S_r P_0\| \leq C r^2.$$

If μ also satisfies (E), then we get the better bound

$$\|P_1 S_r P_0\| \leq C r^{\min\{3, \alpha\}}.$$

Proof. Take $\varphi \in \mathcal{H}_0$, and as in the proof of Lemma 14, write the Taylor expansion:

$$\rho_r(\gamma) \varphi(\xi) = [1 + C_1 \langle \xi, v(\gamma) \rangle + C_2 \langle \xi, v(\gamma) \rangle^2 + O(r^3 |v(\gamma)|^3)] \varphi(\theta_0^{-1} \xi).$$

Similarly to the proof of Lemma 15, we integrate this inequality. To get the first claim, we only need to note that

$$\int \langle \xi, v(\gamma) \rangle d\mu(\gamma) = \left\langle \xi, \int v(\gamma) d\mu(\gamma) \right\rangle = 0$$

because of (C).

Assumption (E) implies that \mathcal{H}_0 consists of even functions, and hence \mathcal{H}_1 contains only odd ones. Since

$$\int \langle \xi, v(\gamma) \rangle^2 d\mu(\gamma) \cdot \varphi(\theta_0^{-1} \xi)$$

is even, it is in the kernel of P_1 . This establishes the second claim. \square

We also need a norm estimate for $P_0 S_r P_0$. As we remarked in (11), this operator is essentially a multiplication operator by the function F . Hence what we need to understand is the behavior of F near the origin.

Lemma 17. *There is a constant C such that*

$$|F(\xi) - (1 - r^2 \Delta(\xi, \xi))| < C r^{\min\{3, \alpha\}},$$

where $\Delta(\xi, \xi)$ is a positive definite quadratic form depending on μ . If $\alpha < 3$, then the above bound can be improved to $o(r^\alpha)$. Furthermore, if μ is symmetric or satisfies (E), then we have the improved bound:

$$|F(\xi) - (1 - r^2 \Delta(\xi, \xi))| < C r^{\min\{4, \alpha\}}.$$

As an immediate corollary, we get that $\|P_0 S_r P_0\| < 1 - cr^2$ for some constant $c > 0$.

Proof. We use the Taylor expansion in the definition of F , i.e. in equation (12) the same way as we did in the previous lemmata:

$$\begin{aligned} F(\xi) = & \int \int 1 + C_1 \langle \sigma \xi, v(\gamma) \rangle - C_2 r^2 \langle \sigma \xi, v(\gamma) \rangle^2 \\ & + C_3 \langle \sigma \xi, v(\gamma) \rangle^3 dm_K(\sigma) d\mu(\gamma) + O(r^{\min\{\alpha, 4\}}), \end{aligned}$$

where C_1, C_3 are constants depending on r and μ , C_2 is an absolute and positive constant, and the implied constant is depends on μ .

First we note that as in the proof of Lemma 16, (C) implies that

$$\int \langle \sigma \xi, v(\gamma) \rangle d\mu(\gamma) = 0$$

for all σ and ξ . Hence the linear term vanishes in the above Taylor expansion.

Second,

$$\Delta(\xi, \xi) := \int \int C_2 r^2 \langle \sigma \xi, v(\gamma) \rangle^2 dm_K(\sigma) d\mu(\gamma)$$

is clearly a K invariant positive semi-definite quadratic form. We only need to show that it is strictly positive definite. Denote by V the maximal subspace of \mathbf{R}^d on which $\Delta(\xi, \xi)$ vanishes. By the definition of Δ , all $v(\gamma)$ is orthogonal to V , which would contradict almost non-degeneracy if $V \neq \{0\}$.

If $\alpha \geq 3$ and (E) is satisfied, then for all ξ , there is $\sigma_\xi \in K$ such that $\sigma_\xi \xi = -\xi$. Then

$$2 \int \langle \sigma \xi, v(\gamma) \rangle^3 dm_K(\sigma) = \int \langle \sigma \xi, v(\gamma) \rangle^3 + \langle \sigma \sigma_\xi \xi, v(\gamma) \rangle^3 dm_K(\sigma) = 0,$$

hence the cubic term of the Taylor expansion of Δ vanishes.

Finally, if μ is symmetric and $\alpha \geq 3$, then

$$2 \int \int \langle \sigma \xi, v(\gamma) \rangle^3 dm_K(\sigma) d\mu(\gamma) = \int \int \langle \xi, \sigma v(\gamma) \rangle^3 + \langle \xi, \sigma \theta(\gamma) v(\gamma^{-1}) \rangle^3 dm_K(\sigma) d\mu(\gamma) = 0.$$

The first equality follows since μ is symmetric and m_K is invariant under multiplication by $\theta(\gamma)$. The second one follows from $\theta(\gamma) v(\gamma^{-1}) = -v(\gamma)$. The proof is finished by combining the above estimates. \square

4.2 Approximating S_r by $P_0 S_r P_0$

Lemma 18. *There are constants c and C , such that the following hold:*

$$\|P_1 \psi'_l\|_2 \leq C(r^2 + e^{-cl}|x_0|r) \quad \text{and} \quad (15)$$

$$\|P_i \psi'_l\|_2 \leq C(r^{\min\{i, \alpha\}} + e^{-cl}|x_0|^i r^i) \quad (16)$$

for $i \in \{2, 3\}$ and $l \geq 0$.

For $l \geq C \log(r^{-1}|x_0| + 2)$, we have

$$\|P_0 \psi'_l\|_2 \leq C e^{-cr^2 l} \quad (17)$$

$$\|P_1 \psi'_l\|_2 \leq C r^2 e^{-cr^2 l} \quad \text{and} \quad (18)$$

$$\|P_i \psi'_l\|_2 \leq C r^{\min\{i, \alpha\}} e^{-cr^2 l} \quad (19)$$

for $i \in \{2, 3\}$.

Moreover, if μ satisfies (E), then we can replace r^2 by $r^{\min\{3, \alpha\}}$ in (15) and (18).

The following proof is very technical although the idea behind it is very simple. The argument is based on induction on l , the norm estimates of the previous section and triangle inequality.

We first give a brief sketch which suggests why the induction works. For simplicity, take $x_0 = 0$ and suppose that the Lemma holds for some l . We can write

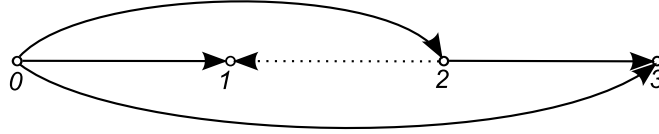
$$P_i \psi'_{l+1} = \sum_{j=0}^3 P_i S_r P_j \psi'_l.$$

We use the induction hypothesis and the Lemmata of the previous section to bound the terms.

What we need is essentially the inequality

$$(1 - \|P_i S_r P_i\|) \|P_i \psi'_l\|_2 > \sum_{j \neq i} \|P_i S_r P_j\| \|P_j \psi'_l\|_2.$$

Notice that if we plug in our estimates that we obtained in the previous section, then all terms on the right hand side is of the same or smaller order of magnitude than the left hand side for any i and j . In the following diagram we draw a directed edge from vertex j to i , if the corresponding term on the right is of the same order as the left hand side.



(The dotted edge is present when (E) is assumed.) Since it does not have a directed cycle, we can set the constants in (17)–(19) following the edges of the diagram such that the induction will work.

Proof. We do not give a separate proof for the last statement, but it will be clear that using the improved estimates in Lemma 16 available under (E), we get the better error terms. Let β_0 and B_0 be constants whose existence is guaranteed by Lemmata 15–17 such that

$$\|P_0 S_r P_0\| \leq e^{-\beta_0 r^2}, \tag{20}$$

$$\|P_i S_r P_i\| \leq e^{-\beta_0} \quad \text{for } i \in \{1, 2, 3\} \tag{21}$$

$$\|P_i S_r P_j\| \leq B_0 r^{\min\{\alpha, |i-j|\}} \quad \text{and} \tag{22}$$

$$\|P_1 S_r P_0\| \leq B_0 r^2. \tag{23}$$

The proof is by induction, and we begin with (15) and (16). We suppose that $|x_0| \geq 1$. In the opposite case the argument is identical; we only need to replace $|x_0|^2$ and $|x_0|^3$ everywhere by $|x_0|$. We show that

$$\begin{aligned}\|P_1\psi'_l\|_2 &\leq C_1(r^2 + e^{-\beta_0 l/2}|x_0|r) \quad \text{and} \\ \|P_i\psi'_l\|_2 &\leq C_i(r^{\min\{i,\alpha\}} + e^{-\beta_0 l/2}|x_0|^i r^i)\end{aligned}$$

for $i \in \{2, 3\}$ and $l \geq 0$, where $C_i > 1$ are suitable constants to be specified later. For $l = 0$, the claim is verified easily by the Taylor expansion of ψ_0 .

Suppose that the claim holds for l , and we prove it for $l + 1$. To estimate

$$\|P_i\psi'_{l+1}\|_2 \leq \sum_{j=0}^3 \|P_i S_r P_j\| \cdot \|P_j\psi'_l\|, \quad (24)$$

we use the induction hypothesis for $\|P_j\psi'_l\|_2$ and the norm estimates (20)–(23). We write for $i = 1$:

$$\begin{aligned}\|P_1\psi'_{l+1}\|_2 &\leq B_0 r^2 + e^{-\beta_0} C_1(r^2 + e^{-\beta_0 l/2}|x_0|r) \\ &\quad + B_0 r C_2(r^2 + e^{-\beta_0 l/2}|x_0|^2 r^2) \\ &\quad + B_0 r^2 C_3(r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2}|x_0|^3 r^3) \\ &\leq ([B_0(1 + r C_2 + r^2 C_3) + e^{-\beta_0} C_1] e^{\beta_0/2}) \cdot (r^2 + e^{-\beta_0(l+1)/2}|x_0|r),\end{aligned}$$

To obtain the last line, we use inequalities of type

$$(r^2 + e^{-\beta_0 l/2}|x_0|^j r^j) \leq e^{\beta_0/2} (r^2 + e^{-\beta_0(l+1)/2}|x_0|^j r^j)$$

and also $r|x_0| < 1$ that we can suppose without loss of generality as we mentioned it at the beginning of Section 4. For $i = 2$:

$$\begin{aligned}\|P_2\psi'_{l+1}\|_2 &\leq B_0 r^2 + B_0 r C_1(r^2 + e^{-\beta_0 l/2}|x_0|r) \\ &\quad + e^{-\beta_0} C_2(r^2 + e^{-\beta_0 l/2}|x_0|^2 r^2) \\ &\quad + B_0 r C_3(r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2}|x_0|^3 r^3) \\ &\leq ([B_0(1 + C_1 + r C_3) + e^{-\beta_0} C_2] e^{\beta_0/2}) \cdot (r^2 + e^{-\beta_0(l+1)/2}|x_0|^2 r^2),\end{aligned}$$

We derive the last line the same way as before, but we also use the inequality $|x_0| \geq 1$. For $i = 3$:

$$\begin{aligned}
\|P_3\psi'_{l+1}\|_2 &\leq B_0 r^{\min\{3,\alpha\}} + B_0 r^2 C_1 (r^2 + e^{-\beta_0 l/2} |x_0| r) \\
&\quad + B_0 r C_2 (r^2 + e^{-\beta_0 l/2} |x_0|^2 r^2) \\
&\quad + e^{-\beta_0/2} C_3 (r^{\min\{3,\alpha\}} + e^{-\beta_0 l/2} |x_0|^3 r^3) \\
&\leq ([B_0(1 + C_1 + C_2) + e^{-\beta_0} C_3] e^{\beta_0/2}) \cdot (r^{\min\{3,\alpha\}} + e^{-\beta_0(l+1)/2} |x_0|^3 r^3).
\end{aligned}$$

Now the claim is satisfied, if we take

$$\begin{aligned}
C_1 &= 2e^{\beta_0/2} B_0 / (1 - e^{-\beta_0/2}) \\
C_2 &= 2e^{\beta_0/2} B_0 (1 + C_1) / (1 - e^{-\beta_0/2}) \\
C_3 &= e^{\beta_0/2} B_0 (1 + C_1 + C_2) / (1 - e^{-\beta_0/2})
\end{aligned}$$

and r is so small that $rC_2 + r^2C_3 < 1$ and $rC_3 < 1$.

The proof of (17)–(19) is very similar. We begin by choosing l_0 such that $e^{-\beta_0 l_0/2} |x_0| r < r^2$. We show by induction that for $l \geq l_0$ the following hold:

$$\begin{aligned}
\|P_0\psi'_l\|_2 &\leq C'_0 e^{-\beta_0 r^2 l/2}, \\
\|P_1\psi'_l\|_2 &\leq C'_1 r^2 e^{-\beta_0 r^2 l/2} \quad \text{and} \\
\|P_i\psi'_l\|_2 &\leq C'_i r^{\min\{i,\alpha\}} e^{-\beta_0 r^2 l/2}
\end{aligned}$$

for $i \in \{2, 3\}$ and $l \geq 0$, where $C'_0 = e^{\beta_0 r^2 l_0/2}$ and for $i > 0$, $C'_i \geq C_i e^{\beta_0 r^2 l_0/2}$ are suitable constants to be specified later. We note that $\beta_0 l_0/2 = \log(|x_0|/r) < r^{-2}$ since $|x_0| r < 1$. Hence $e^{\beta_0 r^2 l_0/2} < e$, so the above constants are bounded independently of x_0 and r .

From the first part of the proof it follows that the claim holds for $l = l_0$. Now we suppose that it holds for a particular $l \geq l_0$ and prove it for $l + 1$. As above, we use (24) along with the induction hypothesis and (20)–(23). We get:

$$\begin{aligned}
\|P_0\psi'_{l+1}\|_2 &\leq e^{-\beta_0 r^2} C'_0 e^{-\beta_0 r^2 l/2} + B_0 r C'_1 r^2 e^{-\beta_0 r^2 l/2} \\
&\quad + B_0 r^2 C'_2 r^2 e^{-\beta_0 r^2 l/2} \\
&\quad + B_0 r^{\min\{3,\alpha\}} C'_3 r^{\min\{3,\alpha\}} e^{-\beta_0 r^2 l/2} \\
&\leq \left([C'_0 e^{-\beta_0 r^2} + B_0 (r^3 C'_1 + r^4 C'_2 + r^{\min\{6,2\alpha\}} C'_3)] e^{\beta_0 r^2/2} \right) \cdot e^{-\beta_0 r^2(l+1)/2},
\end{aligned}$$

$$\begin{aligned}
\|P_1\psi'_{l+1}\|_2 &\leq B_0r^2C'_0e^{-\beta_0r^2l/2} + e^{-\beta_0}C'_1r^2e^{-\beta_0r^2l/2} \\
&\quad + B_0rC'_2r^2e^{-\beta_0r^2l/2} \\
&\quad + B_0r^2C'_3r^{\min\{3,\alpha\}}e^{-\beta_0r^2l/2} \\
&\leq \left([e^{-\beta_0}C'_1 + B_0(C'_0 + rC'_2 + r^{\min\{3,\alpha\}}C'_3)]e^{\beta_0r^2/2}\right) \cdot r^2e^{-\beta_0r^2(l+1)/2},
\end{aligned}$$

$$\begin{aligned}
\|P_2\psi'_{l+1}\|_2 &\leq B_0r^2C'_0e^{-\beta_0r^2l/2} + B_0rC'_1r^2e^{-\beta_0r^2l/2} \\
&\quad + e^{-\beta_0}C'_2r^2e^{-\beta_0r^2l/2} \\
&\quad + B_0rC'_3r^{\min\{3,\alpha\}}e^{-\beta_0r^2l/2} \\
&\leq \left([e^{-\beta_0}C'_2 + B_0(C'_0 + rC'_1 + r^{\min\{2,\alpha-1\}}C'_3)]e^{\beta_0r^2/2}\right) \cdot r^2e^{-\beta_0r^2(l+1)/2},
\end{aligned}$$

$$\begin{aligned}
\|P_3\psi'_{l+1}\|_2 &\leq B_0r^{\min\{3,\alpha\}}C'_0e^{-\beta_0r^2l/2} + B_0r^2C'_1r^2e^{-\beta_0r^2l/2} \\
&\quad + B_0rC'_2r^2e^{-\beta_0r^2l/2} \\
&\quad + e^{-\beta_0}C'_3r^{\min\{3,\alpha\}}e^{-\beta_0r^2l/2} \\
&\leq \left([e^{-\beta_0}C'_3 + B_0(C'_0 + rC'_1 + C'_2)]e^{\beta_0r^2/2}\right) \cdot r^{\min\{3,\alpha\}}e^{-\beta_0r^2(l+1)/2}.
\end{aligned}$$

Take

$$C'_1 \geq \frac{2C'_0B_0e^{\beta_0r^2/2}}{(1 - e^{-\beta_0+\beta_0r^2/2})}, C'_2 \geq \frac{2C'_0B_0e^{\beta_0r^2/2}}{(1 - e^{-\beta_0+\beta_0r^2/2})}, C'_3 \geq \frac{2B_0(C'_0 + C'_2)e^{\beta_0r^2/2}}{(1 - e^{-\beta_0+\beta_0r^2/2})}$$

and observe that the claim holds for $l+1$ if r is sufficiently small. \square

Proof of Lemma 11. By the triangle inequality, we have

$$\|\psi'_l - S_r^l\psi_0\|_2 \leq \sum_{k=0}^{l-1} \|S_r^{l-k-1}(S_r - PS_rP)(PS_rP)^k\psi'_0\|_2 \leq \sum_{k=0}^{l-1} \|(S_r - PS_rP)\psi'_k\|_2$$

To estimate the terms, we write:

$$\|(S_r - PS_rP)\psi'_k\|_2 \leq \sum_{j=0}^3 \|P_\infty S_r P_j\| \cdot \|P_j\psi'_k\|_2,$$

which is valid for $k \geq 1$ since $\psi'_k \in \mathcal{H}_0 \oplus \dots \oplus \mathcal{H}_3$. Recall that P_∞ is the projection to the complement of $\mathcal{H}_0 \oplus \dots \oplus \mathcal{H}_3$.

We use Lemmata 18 and 15. For $1 \leq j \leq C \log(r^{-1}|x_0| + 2)$, we have

$$\|(S_r - PS_rP)\psi'_k\|_2 < C(r^{\min\{4, \alpha\}} + |x_0|^3 r^4 e^{-ck}),$$

while for $k = 0$, we have to add $|x_0|^4 r^4$ to the above estimate. For $k \geq C \log(r^{-1}|x_0| + 2)$, we have

$$\|(S_r - PS_rP)\psi'_k\|_2 < Cr^{\min\{4, \alpha\}} e^{-cr^2 k}.$$

Summing for k , we get the statement of the lemma.

Looking at the proof, it turns out that the constant C in the Lemma ultimately depends on the constant in Lemma 15, hence it can be taken to be arbitrarily small as $r \rightarrow 0$ when $\alpha = 2$. \square

Proof of Lemma 12. By the triangle inequality:

$$\begin{aligned} \|\psi_l - \psi'_l\|_2 &= \|(P_0 S_r P_0)^l \psi_0 - (P S_r P)^l \psi_0\|_2 \\ &\leq \sum_{k=0}^{l-1} \|(P_0 S_r P_0)^{l-k-1} (P_0 S_r P_0 - P S_r P) (P S_r P)^k \psi_0\|_2 \\ &\leq \sum_{k=0}^{l-2} \|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 \\ &\quad + \|(P_0 S_r P_0 - P S_r P) \psi'_{l-1}\|_2. \end{aligned}$$

As in the previous proof, we write

$$\|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 = \sum_{j=1}^3 \|P_0 S_r P_j\| \cdot \|P_j \psi'_k\|_2.$$

Again, we use Lemmata 18, 15, 16 and the estimate $\|P_0 S_r P_0\| \leq 1 - cr^2$ which follows from Lemma 17. For $k \leq \log(r^{-1}|x_0| + 2)$ we can write

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 \leq C e^{-cr^2 l/2} (|x_0| r^2 e^{-ck} + r^3)$$

While for $k \geq \log(r^{-1}|x_0| + 2)$ we get

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 \leq Cr^3 e^{-cr^2(l-1)}$$

and

$$\begin{aligned} \|(P_0 S_r P_0 - P S_r P) \psi'_{l-1}\|_2 &\leq \|(P_0 S_r P_0 - P_0 S_r P) \psi'_{l-1}\|_2 + \|(P_0 S_r P - P S_r P) \psi'_{l-1}\|_2 \\ &\leq Cr^3 e^{-cr^2(l-1)} + Cr^2 e^{-cr^2(l-1)}. \end{aligned}$$

Summing up for k , we get the first statement of the lemma.

If μ is symmetric then S_r is selfadjoint, hence

$$\|P_0 S_r P_1\| = \|P_1 S_r P_0\| \leq C r^2.$$

Using this instead of Lemma 15, we get

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 \leq C r^4 e^{-c r^2 (l-1)},$$

and the better estimate follows after summation.

If μ satisfies (E) instead, then the better estimate in Lemma 18 gives

$$\|P_0 S_r P_0\|^{l-k-1} \cdot \|(P_0 S_r P_0 - P_0 S_r P) \psi'_k\|_2 \leq C r^{\min\{4, \alpha+1\}} e^{-c r^2 (l-1)}.$$

□

4.3 Proof of Lemma 13

Again, we only show the first inequality in the Lemma, the second follows along the same lines by applying the improved estimate in Lemma 17.

Using the Taylor series expansion of ψ_0 together with (C) and finite second moments, we see that

$$\psi_0(\xi) = 1 + O(|x_0|^2 r^2),$$

where the implied constant only depends on μ . Furthermore, we have by (11) that

$$\psi_l(\xi) = \left[\prod_{j=0}^{l-1} F(\theta_0^{-j} \xi) \right] \psi_0(\theta_0^{-l} \xi),$$

where F is given by (12).

Let Δ be the quadratic form appearing in Lemma 17. Define Δ_0 to be its symmetrization by the group generated by θ_0 , i.e.

$$\Delta_0(\xi, \xi) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \Delta(\theta_0^{-i} \xi, \theta_0^{-i} \xi).$$

In light of the previous inequalities it is enough to show that

$$\left| \prod_{j=0}^{l-1} F(\theta_0^{-j} \xi) - e^{-l r^2 \Delta_0(\xi, \xi)} \right| < C e^{-l r^2} r^{\min\{1, \alpha-2\}} \quad (25)$$

for all ξ .

The rest of the proof is devoted to this inequality. By Lemma 17, we have

$$\sum_{j=0}^{l-1} \log F(\theta_0^{-j} \xi) = -r^2 \sum_{j=0}^{l-1} \Delta(\theta_0^{-j} \xi, \theta_0^{-j} \xi) + O(lr^{\min\{3, \alpha\}}).$$

Denote by W the space of quadratic forms on \mathbf{R}^d , and denote by $\Theta_0 \in \text{End}(W)$ the linear transformation induced by θ_0 . It is easily seen that Θ_0 is diagonalizable and all its eigenvalues are on the unit circle of \mathbf{C} . Denote by W_0 the 1-eigenspace of Θ_0 . Hence Δ_0 is the projection of Δ to W_0 . Then on the orthogonal complement of W_0 , we have

$$\sum_{j=0}^{l-1} \Theta_0^{-j} = \frac{\Theta_0^{-l} - 1}{\Theta_0^{-1} - 1}.$$

Thus it follows that

$$\sum_{j=0}^{l-1} \Delta(\theta_0^{-j} \xi, \theta_0^{-j} \xi) = \sum_{j=0}^{l-1} \Theta_0^{-j} \Delta(\xi, \xi) = l \Delta_0(\xi, \xi) + O(1).$$

The implied constant depends on the distance of the non-trivial eigenvalues of Θ_0 to 1.

If we combine our inequalities, we get

$$\sum_{j=0}^{l-1} \log F(\theta_0^{-j} \xi) = -lr^2 \Delta_0(\xi, \xi) + O(r^2 + lr^{\min\{3, \alpha\}}).$$

This immediately implies that there is a constant $c > 0$ such that

$$\prod_{j=0}^{l-1} F(\theta_0^{-j} \xi) < e^{-clr^2}.$$

If we use the inequality $|e^A - e^B| < (A - B) \max\{e^A, e^B\}$, then we get

$$\left| \prod_{j=0}^{l-1} F(\theta_0^{-j} \xi) - e^{-lr^2 \Delta_0(\xi, \xi)} \right| < C e^{-clr^2} (r^2 + lr^{\min\{3, \alpha\}}).$$

To obtain (25) and hence the lemma, we only need to note that

$$e^{-clr^2} l \leq \frac{2}{c} \cdot e^{-clr^2/2} r^{-2}.$$

4.4 Some improvements using (SSR)

The purpose of this section is to give the following slight improvement of the bounds in Proposition 10. The proof depends on the results of Section 3, so it is important to note, that for the argument of Section 3, Proposition 10 itself is enough. In fact, we can even get Theorem 3 without the results of this section at the modest expense of multiplying the first error term by $\log l$.

Proposition 19. *Assume that μ is almost non-degenerate, has finite moments of order $\alpha \geq 2$ and satisfies (C) and (SSR). There are constants C , c and a symmetric positive definite quadratic form $\Delta(\xi, \xi)$ on \mathbf{R}^d invariant under the action of K and θ_0 , such that the following holds*

$$\|S_r^l \psi_0 - e^{-r^2 l \Delta}\|_2 < C(r^{\min\{1, \alpha-2\}} + |x_0|^2 r^2) \cdot (e^{-clr^2} + r^{10d}) \quad (26)$$

for $r < l^{-1/3}$. Moreover, if μ is symmetric or satisfies (E), then we have the better bound:

$$\|S_r^l \psi_0 - e^{-r^2 l \Delta}\|_2 < C(r^{\min\{2, \alpha-2\}} + |x_0|^2 r^2) \cdot (e^{-clr^2} + r^{10d}).$$

Notice that the estimate differs only by the factor $(e^{-clr^2} + r^{10d})$ compared to Proposition 10. We indicate how to modify the argument in the previous sections to obtain this improvement. We only need to sharpen Lemma 11 by the same factor.

First we note, that Proposition 19 gives an improvement only if $l > r^{-2}$ and recall that $l < r^{-3}$ is assumed in the proposition. Hence, we only consider the range $r^{-2} < l < r^{-3}$.

Similarly to the proof of Lemma 11, we write

$$\|S_r^l \psi_0 - (PS_r P)^l \psi_0\|_2 \leq \sum_{j=0}^{l-1} \|S_r^{l-j-1} (S_r - PS_r P) \psi_j'\|_2.$$

To obtain the improvement of Lemma 11, we need to show the following improved estimates for the terms in the above sum:

$$\|S_r^{l-j-1} (S_r - PS_r P) \psi_j'\|_2 \leq C(r^{\min\{4, \alpha\}} + |x_0|^2 r^2 e^{-cj}) \cdot (e^{-clr^2} + r^{20d}). \quad (27)$$

We have already seen that

$$\|(S_r - PS_r P) \psi_j'\|_2 < C(r^{\min\{\alpha, 4\}} + |x_0|^2 r^2 e^{-cj}) e^{-cr^2 j}.$$

We utilize Proposition 4 to estimate the norm of this function when S_r is applied.

If we have

$$\|S_r^m(S_r - PS_rP)\psi'_j\|_2 \leq C(r^{\min\{4,\alpha\}} + |x_0|^2 r^2 e^{-cj}) \cdot r^{20d}$$

for any $m < l - j - 1$, then (27) immediately follows. In the opposite case, we can estimate the Lipschitz norm of

$$\frac{S_r^m(S_r - PS_rP)\psi'_j}{\|S_r^m(S_r - PS_rP)\psi'_j\|_2}$$

by $r^{-20d-4}\sqrt{l}$. (We use here Lemma 5 and the fact that $S_r^l\psi_0$ corresponds to the Fourier transform of the random walk after l steps.) Since $r^2 < \log^{-3}(r^{-20d-4}\sqrt{l})$, we have

$$\|S_r^{m+1}(S_r - PS_rP)\psi'_j\|_2 < e^{-cr^2} \|S_r^m(S_r - PS_rP)\psi'_j\|_2$$

by Proposition 4. Repeated application of this inequality gives the claim (27).

5 The main theorem

We begin this section with discussing the role of assumption (C), that is $\int v(\gamma)d\mu(\gamma) = 0$. We also establish the claim we made in Section 2, which justifies the assumption that K is normalized by θ_0 . We also prove that $\mu^{*(l)}$ is almost non-degenerate for some integer l if μ is non-degenerate. Finally, we deduce Theorem 3 from Propositions 4 and 19.

5.1 Justifying the simplifying assumptions

The purpose of this section is to prove the following simple Lemmata.

Lemma 20. *Assume that there is no point in \mathbf{R}^d except for the origin which is fixed by all elements of K . Then there is a unique point $x \in \mathbf{R}^d$ such that*

$$\int \gamma(x)d\mu(\gamma) = x.$$

The conclusion implies that if we change our coordinate system, and set x to be the origin, then (C) is satisfied.

Proof. Consider the map $\mathbf{R}^d \rightarrow \mathbf{R}^d$:

$$T(x) = \int \gamma(x) d\mu(\gamma) - \int \gamma(0) d\mu(\gamma) = \int \theta(\gamma) x d\mu(\gamma).$$

It is clear that T is a linear transformation.

We show that $x - T(x)$ has trivial kernel. Suppose that $x = T(x)$ for some $x \in \mathbf{R}^d$. Since $|\theta(\gamma)x| = |x|$ for all γ , and $T(x)$ is the average of these points, we must have $\theta(\gamma)x = x$, for μ almost all γ . By our assumption, $x = 0$, hence the kernel of $x - T(x)$ is indeed trivial.

Therefore there is a unique point x such that $x - T(x) = \int \gamma(0) d\mu(\gamma)$, and this is exactly what we needed to prove. \square

Lemma 21. *Let $K < O(d)$ be a compact group, and $\theta_0 \in O(d)$. There is a positive integer l such that θ_0 normalizes the group generated by*

$$K \cup \theta_0^{-1} K \theta_0 \cup \dots \cup \theta_0^{-(l-1)} K \theta_0^{l-1}$$

Proof. It is a well known fact that if K is a compact Lie group, then there is a chain of normal subgroups $K^\circ \triangleleft H \triangleleft K$ such that H/K° is commutative and $[K : H] < C_d$ for a constant C_d depending on d . For a proof in the context of algebraic groups which carries over to compact groups without any changes see [4, Theorem J].

Write K_l for the closure of the group generated by

$$K \cup \theta_0^{-1} K \theta_0 \cup \dots \cup \theta_0^{-(l-1)} K \theta_0^{l-1}$$

and write $K_l^\circ \triangleleft K_l$ for its connected component.

Let $l_0 \geq 1$ be an integer such that $K_l^\circ = K_{l_0}^\circ$ for $l \geq l_0$. (The sequence K_l° stabilizes, since $\dim K_l^\circ$ may grow at most finitely many times.) Assume to the contrary that the sequence K_l does not stabilize and denote by L the closure of their union. Then $K_{l_0}^\circ \triangleleft L$ since $K_{l_0}^\circ$ is normal in all K_l for $l \geq l_0$.

We show that $L^\circ/K_{l_0}^\circ$ is commutative. For any $l \geq l_0$ and $g, h \in K_l$, we have

$$[g^{C_d!}, h^{C_d!}] \in K_{l_0}^\circ$$

hence this property descends to L . Since all elements in a connected compact Lie group are $C_d!$ powers, we have $[L^\circ, L^\circ] < K_{l_0}^\circ$, thus $L^\circ/K_{l_0}^\circ$ is indeed

commutative. Note that L and L° are both normalized by θ_0 which is of crucial importance for what follows.

Write $H_l = K_l \cap L^\circ$. Then clearly $K_l^\circ \triangleleft H_l \triangleleft K_l$, $[K_l : H_l] \leq [L : L^\circ]$ and H_l/K_l° is commutative for $l \geq l_0$. Let $l_1 \geq l_0$ be such that $[K_{l_1} : H_{l_1}] = [L : L^\circ]$ and let g_1, \dots, g_m be a system of representatives for H_{l_1} cosets in K_{l_1} .

We show that $\exp(H_l/K_{l_0}^\circ)$ is constant for $l \geq l_1 + 1$. The *exponent* $\exp(G)$ of a group G is the smallest integer n such that $g^n = 1$ for all $g \in G$. This will be a contradiction since the elements of H_l approximate those of L° and $L^\circ \neq K_{l_0}^\circ$ by our indirect assumption.

Let $l \geq l_1 + 1$. Then all elements of K_{l+1} is of the form

$$g = \prod_{\alpha} \gamma_{\alpha}^{-1} (g_{i_{\alpha}} h_{i_{\alpha}}) \gamma_{\alpha},$$

where $h_{i_{\alpha}} \in H_l$ and $\gamma_{\alpha} \in \{1, \theta_0\}$. For each α , we can write

$$\gamma_{\alpha}^{-1} (g_{i_{\alpha}} h_{i_{\alpha}}) \gamma_{\alpha} = g_{j_{\alpha}} h_{j_{\alpha}} \gamma_{\alpha}^{-1} h_{i_{\alpha}} \gamma_{\alpha},$$

where $g_{j_{\alpha}}$ is the appropriate coset representative and

$$h_{j_{\alpha}} = g_j^{-1} \gamma_{\alpha}^{-1} g_{i_{\alpha}} \gamma_{\alpha} \in H_{l+1} < H_l.$$

We bring all $g_{j_{\alpha}}$ to the left hand side of the product and get that each element of H_{l+1} is of the form

$$h = \prod_{\beta} \gamma_{\beta}^{-1} h_{\beta} \gamma_{\beta},$$

where $h_{\beta} \in H_l$ and $\gamma_{\beta} \in \{1, \theta_0\} \cdot K_{l_1}$. Thus all $\gamma_{\beta}^{-1} h_{\beta} \gamma_{\beta}$ are in L° , in particular they commute modulo $K_{l_0}^\circ$. In addition, the degree of each $h_{\beta} \cdot K_{l_0}^\circ \in H_l/K_{l_0}^\circ$ divides $\exp(H_l/K_{l_0}^\circ)$, hence so is the degree of $h \cdot K_{l_0}^\circ \in H_{l+1}/K_{l_0}^\circ$. This implies that $\exp(H_{l+1}/K_{l_0}^\circ) = \exp(H_l/K_{l_0}^\circ)$ which was to be proved. \square

Lemma 22. *Suppose that μ is non-degenerate. Then there is a positive integer l such that $\mu^{*(l)}$ is almost non-degenerate.*

Proof. By the non-degeneracy assumption, it follows that for each point $x \in \mathbf{R}^d$, there is $l(x)$ such that $\{\gamma(x) : \gamma \in \text{supp}(\mu^{*(l(x))})\}$ is not contained in a proper affine subspace. Indeed, assume to the contrary that this fails, and l_0 and W are such that $\{\gamma(x) : \gamma \in \text{supp}(\mu^{*(l_0)})\}$ spans W and W is of largest

possible dimension. Then $\gamma(x)$ is $d\mu^{*(l)}(\gamma)$ -almost surely contained in $\gamma_0^{l-l_0}W$, where $\gamma_0 \in \text{supp}(\mu)$ is arbitrary. This contradicts to the non-degeneracy of μ .

It is left to show that $l(x)$ is bounded on \mathbf{R}^d . It is easy to see that $\{x : l(x) \leq L\}$ is a Zariski open set for every $L \in \mathbf{Z}$. As $L \rightarrow \infty$ this is an ascending chain which eventually covers \mathbf{R}^d . Therefore the claim follows from the Noetherian property of Zariski open sets. \square

5.2 Proof of Theorem 3

We turn to the proof of the main result of this paper, Theorem 3. Denote by μ the common law of X_i . By assumption, μ has finite moments of order $\alpha > 2$ and satisfies (SSR). Without loss of generality, we can replace μ by $\mu^{*(l_0)}$ for some fixed integer l_0 , hence by Lemmata 22 and 21, we can assume that μ is almost non-degenerate and that K is normalized by θ_0 . Furthermore we assume that μ also satisfies (C) and prove the estimates with $y_0 = 0$. Lemma 20 in the previous section shows that we can reduce the general case of the theorem to this one by changing the coordinate system.

Denote by $\nu_l = \mu^{*(l)}.\delta_{x_0}$ the distribution of the random walk after l steps starting from the point x_0 . To evaluate the left hand sides of the formulae in the statement in Theorem 3, we use Plancherel's formula:

$$\mathbf{E}[f(Y_l)] = \int f(x) d\nu_l = \int \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi.$$

We break the latter integral into two regions. First we consider $|\xi| < l^{-1/3}$ and use Proposition 19 in this region.

Recall from Section 2, that $\text{Res}_r : C(\mathbf{R}^d) \rightarrow C(S^{d-1})$ is the restriction to the sphere of radius r and that

$$\int_{|\xi|=r} |\varphi(\xi)|^2 d\xi = r^{d-1} \|\text{Res}_r \varphi\|_2^2.$$

Recall also that

$$\psi_0(\xi) = \psi_{0,r}(\xi) = e(r\langle x_0, \xi \rangle) = \text{Res}_r(\widehat{\delta_{x_0}})(\xi),$$

and

$$\text{Res}_r \widehat{\nu}_l = S_r^l \psi_{0,r}.$$

For $r \leq l^{-1/3}$, we write:

$$\begin{aligned} \int_{|\xi|=r} \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi &= r^{d-1} \int_{S^{d-1}} [\text{Res}_r \widehat{f}](\xi) \cdot [S_r^l \psi_{0,r}](\xi) d\xi \\ &= \int_{|\xi|=r} \widehat{f}(\xi) e^{-l\Delta(\xi,\xi)} d\xi + r^{d-1} \int_{S^{d-1}} [\text{Res}_r \widehat{f}](\xi) \Psi(\xi) d\xi, \end{aligned} \quad (28)$$

where $\Delta(\xi, \xi)$ is the quadratic form that appears in Proposition 19 and $\Psi = S_r^l \psi_0 - e^{-r^2 l \Delta}$. By Proposition 19,

$$\|\Psi\|_2 \leq C(r^{\min\{1, \alpha-2\}} + |x_0|^2 r^2)(e^{-clr^2} + r^{10d}).$$

Using $|\widehat{f}(\xi)| \leq \|f\|_1$ and the Cauchy-Schwartz inequality, we can bound the second term in (28) by

$$Cr^{d-1}(r^{\min\{1, \alpha-2\}} + |x_0|^2 r^2)(e^{-clr^2} + r^{10d})\|f\|_1.$$

Integrating for $0 \leq r \leq l^{-1/3}$, we can write:

$$\int_{|\xi| \leq l^{-1/3}} \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi = \int_{\xi \in \mathbf{R}^d} \widehat{f}(\xi) e^{-l\Delta(\xi,\xi)} d\xi + O(l^{-\frac{d+\min\{1, \alpha-2\}}{2}} + |x_0|^2 l^{-\frac{d+2}{2}}) \|f\|_1 \quad (29)$$

It is well known that $e^{-l\Delta(\xi,\xi)}$ is the Fourier transform of a Gaussian measure, i.e. there is a quadratic form Δ' and a constant $C_{\Delta'}$ such that

$$\int \widehat{f}(\xi) e^{-l\Delta(\xi,\xi)} d\xi = C_{\Delta'} l^{-d/2} \int f(x) e^{-\Delta'(x,x)/l} dx.$$

Now we see that the first term on the right of (29) is the main term in Theorem 3, while the second one is the first error term. It is also clear that if μ is symmetric or satisfies (E) and we use the improved bounds of Proposition 19, then we get the improved error term claimed in the Theorem.

It is left to show that

$$\int_{|\xi| \geq l^{-1/3}} \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi \leq Ce^{-cl^{1/4}} \|\varphi\|_W, \quad (30)$$

and the proof will be finished.

To this end, we prove a lemma using Proposition 4.

Lemma 23. *Let l be an integer and suppose that $e^{l^{1/4}} > r > l^{-1/3}$ and $|x_0| < e^{l^{1/4}}$. As above, write $\psi_{0,r}(\xi) = e(r\langle x_0, \xi \rangle)$ for $|\xi| = 1$. There is a constant c depending only on μ such that*

$$\|S_r^l \psi_{0,r}\|_2 \leq e^{-cl^{1/4}}.$$

Proof. Choose $1 > c > 0$ to be sufficiently small, to be specified below. Assume to the contrary that the statement is false for some r, l and x_0 . Then for each $j \leq l$, we have

$$\|S_r^j \psi_{0,r}\|_2 > e^{-l^{1/4}},$$

otherwise we get a contradiction from $\|S_r\| \leq 1$. (Recall that $c < 1$.)

To use Proposition 4, we need to estimate the Lipschitz norm of the function $S_r^j \psi_{0,r}$. Recall that

$$S_r^j \psi_{0,r} = \text{Res}_r \widehat{\nu}_j.$$

Using Lemma 5 and the hypothesis on the size of r and x_0 , we get

$$|\frac{d}{d\xi_i} \widehat{\nu}_j(r\xi)| \leq 2\pi r(|x_0| + \int |v(\gamma)| d\mu^{*(j)}(\gamma)) \leq Ce^{2j^{1/4}}$$

for $j \leq l$, with some constant C and for any $1 \leq i \leq d$. Now define

$$\varphi_j = \frac{S_r^j \psi_{0,r}}{\|S_r^j \psi_{0,r}\|_2}.$$

Then we have

$$\|\varphi_j\|_{\text{Lip}} < Ce^{3l^{1/4}}.$$

It follows that

$$r^2 > l^{-2/3} > \log^{-3}(\|\varphi_j\|_{\text{Lip}} + 2).$$

We can apply Proposition 4 for φ_j , and get

$$\|S_r \varphi_j\|_2 \leq e^{-(c'/27)l^{-3/4}},$$

where c' is the constant c from Proposition 4.

If we multiply these inequalities together for $1 \leq j \leq l$, we get

$$\|S_r^l \psi_{0,r}\|_2 \leq e^{-(c'/27)l^{1/4}},$$

a contradiction if we choose c to be less than $c'/27$. □

Similarly as above, we use the Cauchy-Schwartz inequality:

$$\begin{aligned} \int_{|\xi|=r} \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi &= r^{d-1} \int_{S^{d-1}} \text{Res}_r \widehat{f}(\xi) \cdot S_r^l \psi_{0,r}(\xi) d\xi \\ &\leq r^{(d-1)/2} \left(\int_{|\xi|=r} |\widehat{f}(\xi)|^2 d\xi \right)^{1/2} \cdot \|S_r^l \psi_{0,r}\|_2. \end{aligned}$$

We integrate for $r > l^{-1/3}$, and use the Cauchy-Schwartz inequality again:

$$\begin{aligned} \int_{|\xi| \geq l^{-1/3}} \widehat{f}(\xi) \widehat{\nu}(\xi) d\xi &\leq \int_{l^{-1/2}}^{\infty} \left(r^{d+\varepsilon} \int_{|\xi|=r} |\widehat{f}(\xi)|^2 d\xi \right)^{1/2} \cdot (r^{-(1+\varepsilon)/2} \|S_r^l \psi_{0,r}\|_2) dr \\ &\leq \left(\int_{|\xi| \geq l^{-1/3}} |\widehat{f}(\xi)|^2 |\xi|^{d+\varepsilon} d\xi \right)^{1/2} \cdot \left(\int_{l^{-1/3}}^{\infty} \|S_r^l \psi_{0,r}\|_2^2 r^{-1-\varepsilon} dr \right)^{1/2}. \end{aligned}$$

The first integral on the right hand side is bounded by $\|f\|_{W^{2,(d+\varepsilon)/2}}$. By Lemma 23, we have

$$\|S_r^l \psi_{0,r}\|_2^2 r^{-\varepsilon/2} < e^{-cl^{1/4}}$$

for all r with a constant c which depend on ε . Indeed, one can use the lemma for $r \leq e^{-l^{1/4}}$, and simply $\|S_r^l \psi_{0,r}\|_2 \leq 1$ for larger r . This implies (30), and Theorem 3 is proved.

6 Central Limit Theorem

The purpose of this section is to prove Theorem 1. Recall the notation from Sections 1 and 2. We will deduce the theorem from Proposition 10 very similarly to the methods of the previous section.

Notice that the limiting distribution of $Y_l/l^{1/2}$ does not depend on the starting point x_0 . Indeed, let Y'_l be the random walk obtained from the same X_l , but from a different point x'_0 . Since isometries preserve distance, we have

$$|Y_l/\sqrt{l} - Y'_l/\sqrt{l}| = |x_0 - x'_0|/l^{1/2} \rightarrow 0.$$

For the rest of this section, take $x_0 = 0$.

Denote by $W \subset \mathbf{R}^d$ the linear subspace of vectors fixed by K . By Lemma 20, we can choose the origin in such a way, that

$$v_0 := \mathbf{E}(Y_1) = \int \gamma(0) d\mu(\gamma) \in W.$$

Define the random isometries X'_l by $X'_l(x) = X_l(x) - v_0$. Since $v_0 \in W$, we have

$$Y'_l := X'_l(X'_{l-1}(\dots(0))) = Y_l - lv_0. \quad (31)$$

Denote by μ' the law of the random isometry X'_1 . Then μ' satisfies (C). In what follows, we assume that $\mu = \mu'$, and prove the theorem with $v_0 = 0$. In light of (31), this implies the theorem without the assumption $\mu = \mu'$ as well.

Let $\nu_l = \mu^{*(l)} \cdot \delta_0$ be the law of Y_l . Denote by $\psi_0 \in L^2(S^{d-1})$ the constant function $\psi_0(\xi) = 1$. Similarly to Section 5.2, we can write

$$\text{Res}_r \widehat{\nu}_l(\xi) = S_r^l \psi_0(\xi).$$

Fix an arbitrary constant $R > 0$. Let Δ be the quadratic form from Proposition 10. Let λ be the Gaussian measure with Fourier transform

$$\widehat{\lambda}(\xi) = e^{-\Delta(\xi, \xi)}.$$

Proposition 10 implies

$$\|\text{Res}_{r/\sqrt{l}} \widehat{\nu}_l - \text{Res}_r \widehat{\lambda}\|_2 \rightarrow 0$$

as $l \rightarrow \infty$, uniformly for $r < R$. Let f be a function, such that $\widehat{f}(\xi) = 0$ for $|\xi| > R$. Then by Plancherel and Cauchy-Schwartz:

$$\begin{aligned} \left| \int f(x/\sqrt{l}) d\nu_l(x) - \int f(x) d\lambda(x) \right| &= \left| \int \widehat{f}(\xi) (\widehat{\nu}_l(\xi/\sqrt{l}) - \widehat{\lambda}(\xi)) d\xi \right| \\ &\leq \|f\|_1 \cdot \int_{|\xi| < R} |\widehat{\nu}_l(\xi/\sqrt{l}) - \widehat{\lambda}(\xi)| d\xi \\ &\leq \|f\|_1 \cdot C_{R,d} \int_0^R \|\text{Res}_{r/\sqrt{l}} \nu_l - \text{Res}_r \widehat{\lambda}\|_2 dr \rightarrow 0, \end{aligned}$$

where $C_{R,d}$ is a constant depending on R and d .

Since we can approximate any continuous function by those which have compactly supported Fourier transform, the proof is complete.

7 Local Limit Theorem

We finish the paper with the proof of Theorem 2. Since we do not assume that μ satisfies (SSR), we need to find a suitable replacement for Proposition

4. In this case, our estimate is weaker; we are no longer able to control the dependence of the constants on r for large values of r . We give these estimates in the next two sections. Using these, we will still be able to conclude the local limit theorem but only on scales $O(1)$ in contrast to Theorem 3.

We introduce some notation. Recall that G is the closure of the group generated by $\text{supp}(\tilde{\mu} * \mu)$ and $\text{supp}(\mu)$ is contained in the coset $\gamma_0 G$. We denote by K the closure of $\theta(G)$ and by K° its connected component. By Lemma 21, we can assume that K is normalized by $\theta(\gamma_0)$.

It is easy to see, that we can decompose \mathbf{R}^d as an orthogonal sum of subspaces $V_{ss} \oplus V_a \oplus V_o$, such that action of K° is semi-simple on V_{ss} , Abelian on V_a and trivial on V_o . Since K° is invariant under conjugation by elements of K and $\theta(\gamma_0)$, it follows that these subspaces are invariant under K and $\theta(\gamma_0)$, as well. We denote by S^i the unit sphere in V_i , where i is either ss , a or o , furthermore we denote by $\pi_i(\gamma) \in \text{Isom}(V_i)$ the restriction of $\gamma \in \text{Isom}(\mathbf{R}^d)$ to the subspace V_i . In addition, we will denote by $\pi_i(\mu)$ the probability measure on $\text{Isom}(V_i)$ which is the pushforward of μ under π_i . Finally, θ_i and v_i (resp.) are shorthands for $\pi_i \circ \theta$ and $\pi_i \circ v$ (resp.).

In Section 7.1 we generalize Proposition 4 to $V_{ss} \oplus V_a$. This will give us a useful estimate on $\widehat{\nu}_l(\xi)$ roughly in the range $|\pi_{ss}\xi|^2 + |\pi_a\xi|^2 > 1/l$. In Section 7.2 we give an estimate that is useful in the opposite case. Finally in Section 7.3 we combine them to conclude Theorem 2.

7.1 Estimates for large $\xi_{ss} + \xi_a$

Fix three non-negative real parameters r_{ss}, r_a and r_o . Analogously to ρ_r , we introduce the unitary representation ρ_{r_{ss}, r_a, r_o} of the group $\text{Isom}(\mathbf{R}^d)$ on the space $L^2(S^{ss} \times S^a \times S^o)$ via the following formula:

$$\begin{aligned} \rho_{r_{ss}, r_a, r_o}(\gamma) \varphi(\xi_{ss}, \xi_a, \xi_o) &= e(r_{ss} \langle \xi_{ss}, v_{ss}(\gamma) \rangle + r_a \langle \xi_a, v_a(\gamma) \rangle + r_o \langle \xi_o, v_o(\gamma) \rangle) \\ &\quad \cdot \varphi(\theta_{ss}(\gamma)^{-1} \xi_{ss}, \theta_a(\gamma)^{-1} \xi_a, \theta_o(\gamma)^{-1} \xi_o). \end{aligned}$$

Here $\varphi \in L^2(S^{ss} \times S^a \times S^o)$, $\xi_{ss} \in S^{ss}$, $\xi_a \in S^a$ and $\xi_o \in S^o$. This representation corresponds to the action of γ on the Fourier transform of a measure restricted to the product of the spheres of radii r_{ss}, r_a, r_o (resp.) in V_{ss}, V_a, V_o (resp.).

The purpose of this section is to study the operator

$$S_{r_{ss}, r_a, r_o} = \int \rho_{r_{ss}, r_a, r_o}(\gamma) d\mu(\gamma)$$

acting on $L^2(S^{ss} \times S^a \times S^o)$

Proposition 24. *Suppose that μ is almost non-degenerate, has finite moments of order 2 and satisfies (C). Then for every $R > 0$, there is a constant $c > 0$ depending only on μ and R such that the following hold.*

Let $R \geq r_{ss}, r_a, r_o \geq 0$. Let $\varphi \in \text{Lip}(S^{ss} \times S^a \times S^o)$ with $\|\varphi\|_2 = 1$. Then

$$\|S_{r_{ss}, r_a, r_o} \varphi\|_2 \leq 1 - c \min\{r_{ss}^2 + r_a^2, \frac{1}{\log^3(\|\varphi\|_{\text{Lip}} + 2)}\}.$$

For the rest of this section, we fix r_{ss}, r_a, r_o and omit them from the indices everywhere to save a considerable amount of ink.

Our first goal is to symmetrize μ and replace it with a measure μ_1 such that $\text{supp}(\mu_1) \subset K^\circ$.

Lemma 25. *We can write $(\tilde{\mu} * \mu)^{(L)} = p\mu_1 + q\mu_2$, with $p, q > 0$, where μ_1 and μ_2 are probability measures on $\text{Isom}(\mathbf{R}^d)$ and $L \geq 1$ is an integer depending on μ . Furthermore, μ_1 is almost non-degenerate, symmetric, has finite moments of order 2, and the closure of the group generated by $\text{supp}(\mu_1)$ is K° .*

Proof. Fix an integer L and write

$$G^\circ = \{\gamma \in G : \theta(\gamma) \in K^\circ\},$$

let $p = (\tilde{\mu} * \mu)^{(L)}(G^\circ)$ and let μ_1 be $1/p$ times the restriction of $(\tilde{\mu} * \mu)^{(L)}$ to G° . The only non-trivial property to check is that almost non-degeneracy holds if L is sufficiently large. It is enough to check this condition for one point x , because then the claim follows from the same Noetherian property argument as in Lemma 22.

Denote by o the order of K/K° . Using the central limit theorem for the measure $\tilde{\mu} * \mu$, we can find an integer L_0 , and a finite set

$$A \subset \{\gamma(x) : \gamma \in \text{supp}((\tilde{\mu} * \mu)^{(L_0)})\}$$

which approximates an $(o+1) \times \cdots \times (o+1)$ grid. The approximation can be arbitrarily good, if L_0 is sufficiently large. All that we need is that a proper affine subspace intersects A in at most $|A|/(o+1)$ points.

Then by the pigeon hole principle, there is $\theta_1 \in K$ such that

$$B := \{\gamma(x) : \gamma \in \text{supp}((\tilde{\mu} * \mu)^{(L_0)}), \theta(\gamma) \in \theta_1 K^\circ\}$$

is not contained in a proper affine subspace. Then the claim follows for $L = 2L_0$. Take any $\gamma_1 \in \text{supp}((\tilde{\mu} * \mu)^{(L_0)})$ with $\theta(\gamma_1) \in \theta_1 K^\circ$ and observe that $\gamma_1^{-1}(B)$ is in the set of images of x under elements of $\text{supp}(\mu_1)$. \square

For the rest of the proof we work with μ_1 and assume that it satisfies the properties claimed in Lemma 25. Moreover, we assume that μ_1 has property (C) which is justified by Lemma 20 after changing the origin.

We define the operator

$$T\varphi = \int \rho(\gamma)\varphi d\mu_1(\gamma)$$

on $L^2(S^{ss} \times S^a \times S^o)$, and set out to prove for it an inequality analogous to the one claimed in Proposition 24.

We fix $\varphi \in C^1(S^{ss} \times S^a \times S^o)$. As in Section 4, the heart of the proof is the study of the set

$$B(\varepsilon) := \{\gamma \in \text{Isom}(\mathbf{R}^d) : \|\rho(\gamma)\varphi - \varphi\|_2 < \varepsilon\}.$$

The next two Lemmata is obtained by a simple variation on the arguments in Section 4.

Lemma 26. *Let $\varepsilon > 0$ and let $l_1 = C_1(r_{ss}^{-2} + \log^3(\|\varphi\|_{\text{Lip}} + 2))$, where C_1 is a suitably large constant depending on μ and ε . Suppose that $\mu_1^{*(l_1)}(B(\varepsilon)) > 9/10$. Then there is a set $X \subset B(64\varepsilon)$ such that*

$$\theta_{ss}(X) = \pi_{ss}(K^\circ) \quad \text{and} \quad \pi_a(X) = \{1\} = \pi_o(X).$$

Proof. Following the proof of Lemma 6, it is easy to find a subset $X_0 \subset B(4\varepsilon)$ such that $\pi_{ss}(\theta(X_0)) = \pi_{ss}(K)$. Consider $X_1 = [X_0, X_0]$ and $X = [X_1, X_1]$. Clearly $\pi_o(X_1) = \{1\}$, and $\pi_a(X_1)$ consists of translations. Therefore $\pi_a(X) = \{1\} = \pi_o(X)$, and $X \subset B(64\varepsilon)$ follows from the triangle inequality. Since every element is a commutator in a connected semi-simple compact Lie group (see [12]), we have $\pi_{ss}(X) = \pi_{ss}(K^\circ)$ which finishes the proof. \square

Lemma 27. *Let $\varepsilon > 0$ be arbitrary and l_1 and C_1 be as in the previous lemma. Let $l_2 = C_2(r_{ss}^{-2} + \log^3(\|\varphi\|_{\text{Lip}} + 2))$, where C_2 is a suitably large constant depending on μ_1 and C_1 . Suppose that $\mu^{*(l_i)}(B(\varepsilon)) > 9/10$, for $i = 1, 2$.*

Then there is a constant $c > 0$ depending on μ_1 such that the following hold. For any unit vector $u_0 \in S^{ss}$, there is an element $\gamma'_1 \in B(386\varepsilon)$ such that

$$\begin{aligned} \theta_{ss}(\gamma'_1) = 1, \quad |v_{ss}(\gamma'_1)| < r_{ss}^{-1}/2, \quad \langle v_{ss}(\gamma'_1), u_0 \rangle > cr_{ss}^{-1}, \quad \text{and} \\ \pi_a(\gamma'_1) = 1 = \pi_0(\gamma'_1). \end{aligned}$$

Proof. Consider the projection to V_{ss} and repeat the argument in Lemmata 7–9, except that instead of the set X constructed in Lemma 6, use the one constructed in Lemma 26. We need to use six elements of X , and since now they are in $B(64\varepsilon)$ instead of $B(4\varepsilon)$, the resulting element γ'_1 will be in $B(386\varepsilon)$. Recall from the proof of Lemma 9 that γ'_1 is of the form $g_1\gamma_1g_1^{-1}g_2\gamma_1^{-1}g_2^{-1}$, where $g_1, g_2 \in X$. Since $\pi_a(g_1) = \pi_a(g_2) = 1$, we have $\pi_a(\gamma'_1) = \pi_a(\gamma_1)\pi_a(\gamma_1^{-1}) = 1$, and a similar calculation applies to the projection to V_o . This finishes the proof. \square

In the above lemma we constructed a translation in V_{ss} . The next goal will be to construct a translation in V_a . This is done in the next two lemmata by adapting the method of Guivarc'h [13]. Denote by G_1 the closure of the group generated by $\text{supp}(\mu_1)$.

Lemma 28. $\pi_a([G_1, G_1])$ is the additive group of the vectorspace V_a .

Here $[G_1, G_1]$ denotes the derived subgroup of G_1 , not just the set of commutators.

Proof. Clearly $H := \pi_a([G_1, G_1])$ is a subgroup of the additive group of V_a , and it is invariant under the action of $\pi_a(K^\circ)$. Since $\pi_a(K^\circ)$ is connected, every connected component of H is invariant under $\pi_a(K^\circ)$. Every such component is an affine subspace of V_a . The point of such an affine subspace which is closest to the origin is a fixed point of $\pi_a(K^\circ)$. By the definition of V_a , the only fixed point is the origin. Therefore it follows that H is a linear subspace of V_a invariant under the action of $\pi_a(K^\circ)$.

Assume to the contrary that H is a proper subspace of V_a . Let W be a two dimensional subspace of V_a which is invariant under $\pi_a(K)$ and orthogonal to H . By projecting the translation part to W , G_1 naturally embeds to $\text{Isom}(W)$; denote by G_W the image. Then G_W is commutative (since it has a trivial commutator) but has a non-trivial rotation part (since $\pi_a(K^\circ)$ acts on W non-trivially), hence it consists of rotations around the same point $x \in W$. This means that μ_1 almost every image of x is orthogonal to W , a contradiction to almost non-degeneracy. \square

Lemma 29. *Let ε , l_1 and C_1 be as in Lemma 26. Suppose that $\mu_1^{*(l_1)}(B(\varepsilon)) > 9/10$. Then for every $u_0 \in V_a$, there are $c > 0$, $v \in V_a$ with $|v - u_0| < |u_0|/10$ and an integer L such that the following holds. Let M be an arbitrary positive integer and assume that $\mu_1^{*(2)}(B(\varepsilon/M)) > 1 - c$. Then there is $\gamma'_1 \in B(L\varepsilon)$ such that*

$$v_a(\gamma'_1) = Mv, \quad \theta_a(\gamma'_1) = 1 \quad \text{and} \quad \pi_{ss}(\gamma'_1) = 1 = \pi_o(\gamma'_1).$$

The vector v may also depend on φ , but c and L depend only on μ and u_0

To motivate the Lemma, we indicate how it is applied. We set the parameters in such a way that Mu_0 is of length approximately $1/r_a$ which will lead us to a contradiction. When $r_a \leq 1$, we take u_0 to be of length ≈ 1 and $M \approx 1/r_a$. When $r_a \geq 1$, we take u_0 to be of length $\approx 1/r_a$ and $M = 1$. We also add, that when the condition $\mu_1^{*(2)}(B(\varepsilon/M)) > 1 - c$ fails, we can prove $\|T\varphi\|_2 \leq 1 - c/M^2$, which is exactly our aim with the above choices of parameters.

Observe that the constants c and L depend on u_0 in an uncontrolled way. For estimating the Fourier transform of the random walk in a fixed ball, it will be enough to consider a finite number of points u_0 . However we are no longer able to control the behavior in growing balls.

Proof. There is $\gamma_2 \in G_1$, such that $\theta_a(\gamma_2)$ does not have any fixed vectors in V_a except for 0. This is an open condition, so we can assume that $\gamma_2 \in \text{supp } \mu_1^{*(m)}$ for some integer m depending on μ . Thus $\gamma_2 = g_1 \cdots g_m$ for some $g_i \in \text{supp } \mu_1$.

We can find a small ball U_i around each g_i such that $\theta_a(g'_1 \cdots g'_m)$ does not have any fixed vectors in V_a except for 0, for any choice of $g'_i \in U_i$. We set the constant c in the lemma so that $\mu_1^{*(2)}(U_i) > c$ for all i . This allows us to find an element

$$\gamma'_2 = g'_1 \cdots g'_m \in B(m\varepsilon/M)$$

such that $\theta_a(\gamma'_2)$ does not have any fixed vectors in V_a except for 0.

Then there is a vector $u_1 \in V_a$ such that $u_0 = u_1 - \theta_a(\gamma'_2)u_1$. (Since $\theta_a(\gamma'_2)$ has no fixed vectors, $1 - \theta_a(\gamma'_2)$ has trivial kernel.) For reasons that will be clear at the end of the proof we try to find an element $\gamma \in B(L\varepsilon/M)$ such that $v_a(\gamma)$ approximates u_1 instead of u_0 which is the objective in the lemma.

By Lemma 28, we have $u_1 \in \pi_a([G_1, G_1]) \subset \pi_a(G_1)$. Using an argument very similar to the one above, we can find a vector v_1 and an element $\gamma'_3 \in$

$B(m\varepsilon/M)$ (with a larger m perhaps) such that $v_a(\gamma'_3) = v_1$, $|v_1 - u_1| < |u_0|/20$ and $\theta(\gamma'_3) = 1$.

Now we make use of the set X constructed in Lemma 26 to cancel the rotation parts of γ'_2 and γ'_3 in the V_{ss} component. Let $h_2, h_3 \in X$ be such that $\theta_{ss}(h_2) = \theta_{ss}(\gamma'_2)^{-1}$ and $\theta_{ss}(h_3) = \theta_{ss}(\gamma'_3)^{-M}$. Then $h_2 \cdot \gamma'_2$ and $h_3 \cdot (\gamma'_3)^M$ act on $V_{ss} \oplus V_o$ by translation, hence

$$\gamma'_1 := [h_3 \cdot (\gamma'_3)^M, h_2 \gamma'_2]$$

acts trivially on V_{ss} and V_o . On the other hand, an easy calculation shows that

$$\pi_a(\gamma'_1) = (M(v_1 - \theta_a(\gamma'_2)v_1), 1)$$

and $\gamma'_1 \in B(4(m+64)\varepsilon)$ which was to be proved. \square

Proposition 24. Recall the definition of the operator T from the beginning of Section 7.1. Without any significant changes to the argument in the proof of Proposition 4, we can deduce from Lemma 27 the estimate

$$\|T\varphi\|_2 \leq 1 - c' \min\{r_{ss}^2, \frac{1}{\log^3(\|\varphi\|_{\text{Lip}} + 2)}\}.$$

We suppress the details but carry out a similar argument which proves

$$\|T\varphi\|_2 \leq 1 - c' \min\{r_a^2, \frac{1}{\log^3(\|\varphi\|_{\text{Lip}} + 2)}\}. \quad (32)$$

There is a unit vector $u \in S^a$ such that

$$\int_{\xi_{ss} \in S^{ss}, \xi_o \in S^o} \int_{\xi_a \in S^a: |\xi_a - u| < 1/10} |\varphi(r_{ss}\xi_{ss}, r_a\xi_a, r_o\xi_o)|^2 d\xi_{ss} d\xi_a d\xi_o > \varepsilon_0^2 \quad (33)$$

for a constant ε_0 which depends only on the dimension of V_a . Moreover, we can choose u from a fixed finite sufficiently dense subset of S^a . If $r_a < 1$, let $M = \lfloor 10r_a^{-1} \rfloor$ and let $M = 1$ otherwise. If $r_a < 1$, then let $u_0 = u/20$, otherwise let $u_0 = 5u/\lceil 10r_a \rceil$. Let C_1 , c and L be the constants from Lemma 29 with this choice of u_0 . (Note that the possible values for u_0 are in a finite set which depends only on the dimension of V_a and R .) Set $\varepsilon = \varepsilon_0/L$.

Assume to the contrary that $\mu_1^{*(l_1)}(B(\varepsilon)) > 9/10$ and $\mu^{*(2)}(B(\varepsilon/M)) > 1 - c$. Then we can apply Lemma 29. Let $v \in V_a$ and γ'_1 be as in the Lemma.

Then

$$\begin{aligned}\varepsilon_0^2 &= L^2 \varepsilon^2 \geq \|\rho(\gamma_1')\varphi - \varphi\|_2^2 \\ &\geq \int_{\xi_{ss} \in S^{ss}, \xi_o \in S^o} \int_{|\xi_a - u| < 1/10} |1 - e(\langle Mv, r_a \xi_a \rangle)|^2 |\varphi(r_{ss}\xi_{ss}, r_a \xi_a, r_o \xi_o)|^2 d\xi_{ss} d\xi_a d\xi_o\end{aligned}$$

With the above definitions, Mu_0 and Mv are very close to $r_a^{-1}u/2$, hence $e(\langle Mv, \xi_a \rangle)$ is very close to -1 in the domain of integration. In particular, $|1 - e(\langle Mv, \xi_a \rangle)| > 1$ which is in contradiction to (33).

Now there are two possibilities: Either $\mu_1^{*(l)}(B(\varepsilon)) \leq 9/10$, which implies (32) as we have seen in the proof of Proposition 4. Or else $\mu_1^{*(2)}(B(\varepsilon/M)) \leq 1 - c$. If $\gamma_1^{-1} \cdot \gamma_2 \in B(\varepsilon/M)$, then

$$\operatorname{Re}(\langle \rho(\gamma_2)\varphi, \rho(\gamma_1)\varphi \rangle) = \operatorname{Re}(\langle \rho(\gamma_1^{-1} \cdot \gamma_2)\varphi, \varphi \rangle) \leq 1 - \varepsilon^2/2M^2.$$

Hence (32) follows:

$$\|T\varphi\|_2^2 = \int \operatorname{Re}(\langle \rho(\gamma_2)\varphi, \rho(\gamma_1)\varphi \rangle) d\mu(\gamma_1) d\mu(\gamma_2) \leq 1 - c\varepsilon^2/2M^2.$$

If p and q are as in Lemma 25 and L' is the number L from that lemma, then we can conclude

$$\|(S_{r_{ss}, r_a, r_o}^* S_{r_{ss}, r_a, r_o})^{L'} \varphi\|_2 \leq p \|T\varphi\|_2 + q$$

which in turn implies the proposition. □

7.2 Estimates for small $\xi_{ss} + \xi_a$

In this section we employ the methods of Section 4 (i.e. Taylor expansion) to establish some estimates that we will use in the range where ξ_{ss} and ξ_a are small.

We continue to use the notation, V_{ss}, V_a, V_0 , etc. introduced in the beginning of Section 7.

We fix a couple of parameters $r_{ss}, r_a, R, \delta > 0$ and an integer α . R will be the radius of the ball in V on which we bound the Fourier transform of the random walk, α controls the number of terms in our Taylor expansion and δ is a parameter which control the distance of the ξ_o coordinate from special points, where certain singularities occur, to be discussed below.

It will be helpful for us to work on slightly different function spaces than before. In the course of the proof we will find a finite set $Y_\alpha \subset V_o$ which contains 0 and invariant under $\theta_o(\text{supp } \mu)$. We will work on the following two function spaces.

$$\begin{aligned}\mathcal{H}_1 &:= L^2(S^{ss} \times S^a \times \{\xi_o \in V_o : R^{-1} \leq |\xi_o| \leq R, \text{dist}(\xi_o, Y_\alpha) \leq \delta\}) \\ \mathcal{H}_2 &:= L^2(S^{ss} \times S^a \times \{\xi_o \in V_o : |\xi_o| \leq R, \text{dist}(\xi_o, Y_\alpha) > \delta\}).\end{aligned}$$

We denote the orthogonal projections to these subspaces by π_1 and π_2 respectively. We assume without loss of generality that R is sufficiently big so that Y_α does not contain a point smaller than $R^{-1} + \delta$, hence the above two spaces cover the ball of radius R in V_o apart from the ball of radius δ around the origin.

We work with the operator

$$S_{r_{ss}, r_a} \varphi = \int \rho_{r_{ss}, r_a}(\gamma) \varphi d\mu(\gamma), \quad \text{where}$$

$$\rho_{r_{ss}, r_a}(\gamma) \varphi(\xi_{ss}, \xi_a, \xi_o) = e(\langle r_{ss} \xi_{ss} + r_a \xi_a + \xi_o, v(\gamma) \rangle) \varphi(\theta(\gamma)^{-1}(\xi_{ss}, \xi_a, \xi_o)).$$

This operator acts on both \mathcal{H}_1 and \mathcal{H}_2 .

As before, let x_0 be the starting point of the random walk, and $\psi_0 \in L^2(S^{ss} \times S^a, V_o)$ defined by

$$\psi_0(\xi_{ss}, \xi_a, \xi_o) = e(\langle x_0, r_{ss} \xi_{ss} + r_a \xi_a + \xi_o \rangle),$$

i.e. the restriction of $\widehat{\delta}_{x_0}$ to the product of spheres of radius r_{ss}, r_a and V_o . Finally, put $\psi_i = \pi_i(\psi_0)$ for $i = 1, 2$.

Proposition 30. *Assume that μ is non-degenerate and has finite moments of order 2. Then there is a number C depending on μ and R such that*

$$\|S_{r_{ss}, r_a}^l \psi_1\|_{\mathcal{H}_1} < C \log(s^{-1}) s \delta^{\dim V_o / 2}$$

if s is a number which satisfies $l > C \log(s^{-1})$, $s > r_{ss} + r_a$ and δ satisfies (34) with some C_1 depending on μ and R .

Proposition 31. *Assume that μ is non-degenerate and has finite moments of order α for some $\alpha \geq 2$. Then there is a number C depending on μ, R and α such that*

$$\|S_{r_{ss}, r_a}^l \psi_2\|_{\mathcal{H}_2} < C \log(s^{-1})^{\alpha/2+1} s^\alpha \delta^{-\alpha-2}$$

if s is number which satisfies $l \geq C \log(s^{-1}) \delta^{-2}$, $s > r_{ss} + r_a$ and δ satisfies (34) with some C_1 depending on μ and R .

We indicate the approximate values of the parameters that we will set in the next section. We take $s = C(\log^{1/2} l)l^{-1/2}$ and use the above two propositions for $r_{ss} + r_a < s$ while in the opposite region we use the result of the previous section. We put $\delta = l^{-\beta}$, where β is slightly smaller than $1/2$ so that if we integrate the estimate in Proposition 30 for r_{ss} and r_a , then we get a contribution which is $o(l^{-d/2})$. Then we take α sufficiently large so that $(s/\delta)^\alpha$ will be a suitably large negative power of l . The details of these calculations are carried out in the next section.

With these choices of the parameters, we see that δ satisfies the inequalities

$$C_1^{-1} \geq \delta \geq C_1(r_{ss} + r_a) \quad (34)$$

for arbitrarily large C_1 if l is sufficiently large. From now on, we assume that this inequality is satisfied with a suitably large C_1 depending on μ and R .

We will also use the measure μ_1 constructed in Lemma 25, and assume that μ_1 also satisfies (C). In fact, we will need the almost non-degeneracy property for μ_1 in a slightly stronger form. We require that

$$\{v_o(\gamma) : \gamma \in \text{supp } \mu_1, |v_o(\gamma)| < 1/(2R)\} \quad (35)$$

spans V_o . This property can be obtained by a simple variation of the argument in Lemma 25. First we note that (35) does not depend on the choice of the origin because K° acts trivially on V_o . This means that we do not need the Noetherian property argument; which is good news since there is no reason why a condition like (35) would be Zariski open. The only change we need to make to the proof of Lemma 25 is that we demand that the grid A is inside a ball of diameter $1/(2R)$. We can achieve this by using the non-degeneracy property for μ instead of the central limit theorem.

Moreover, we work with the operator T introduced in the previous section, but on a different function space: We disintegrate it with respect to the ξ_o variable, i.e. we look at the following operator acting on the space $L^2(S^{ss} \times S^a)$:

$$T_{r_{ss}, r_a, \xi_o} \varphi = \int \rho_{r_{ss}, r_a, \xi_o}(\gamma) \varphi d\mu_1(\gamma), \quad \text{where}$$

$$\rho_{r_{ss}, r_a, \xi_o}(\gamma) \varphi(\xi_{ss}, \xi_a) = e(\langle r_{ss} \xi_{ss} + r_a \xi_a + \xi_o, v(\gamma) \rangle) \varphi(\theta_{ss}^{-1}(\gamma) \xi_{ss}, \theta_a^{-1}(\gamma) \xi_a).$$

Note that isometries in $\text{supp } (\mu_1)$ have trivial rotation component on V_o .

Lemma 32. *There is a constant C depending on μ_1 such that*

$$\|T_{r_{ss}, r_a, \xi_o} - T_{0,0,\xi_o}\| < C(r_{ss} + r_a).$$

Proof. Let $\varphi \in L^2(S^{ss} \times S^a)$. Then by Taylor's theorem

$$\begin{aligned} T_{r_{ss}, r_a, \xi_o} \varphi(\xi_{ss}, \xi_a) &= \int (1 + O(r_{ss} + r_a)) e(\langle \xi_o, v_o(\gamma) \rangle) \\ &\quad \times \varphi(\theta_{ss}(\gamma)^{-1} \xi_{ss}, \theta_a(\gamma)^{-1} \xi_a) d\mu_1(\gamma) \\ &= (1 + O(r_{ss} + r_a)) \cdot T_{0,0,\xi_o} \varphi(\xi_{ss}, \xi_a). \end{aligned}$$

In the second equality we used that μ_1 has finite moments of order 1. \square

Lemma 33. *There are constants c and C which depend only on μ_1 and R such that the following holds. Suppose that $\varphi \in L^2(S^{ss} \times S^a)$, and there is $|\xi_o| \leq R$ such that $T_{0,0,\xi_o} \varphi = \varphi$. Then*

$$\|T_{r_{ss}, r_a, \xi'_o} \varphi\|_2 < 1 - c|\xi_o - \xi'_o|^2 + C(r_{ss}^2 + r_a^2)$$

for every $r_{ss}, r_a \geq 0$ and $\xi'_o \in V_o$ with $|\xi'_o| < R$.

Proof. Since $T_{0,0,\xi_o}$ is an average of unitary operators, we must have $\rho_{0,0,\xi_o}(\gamma)\varphi = \varphi$ for all $\gamma \in \text{supp}(\mu_1)$. Since μ_1 has finite second moments,

$$\begin{aligned} T_{r_{ss}, r_a, \xi'_o} \varphi &= \int \rho_{r_{ss}, r_a, \xi'_o}(\gamma) \rho_{0,0,\xi_o}(\gamma^{-1}) \varphi d\mu_1(\gamma) \\ &= \varphi \cdot \int e(r_{ss} \langle \xi_{ss}, v_{ss}(\gamma) \rangle + r_a \langle \xi_a, v_a(\gamma) \rangle + \langle \xi'_o - \xi_o, v_o(\gamma) \rangle) d\mu_1(\gamma) \\ &= \varphi \cdot \left[\int (1 - 2\pi i r_{ss} \langle \xi_{ss}, v_{ss}(\gamma) \rangle - 2\pi i r_a \langle \xi_a, v_a(\gamma) \rangle) \right. \\ &\quad \left. \times e(\langle \xi'_o - \xi_o, v_o(\gamma) \rangle) d\mu_1(\gamma) + O(r_a^2 + r_{ss}^2) \right]. \end{aligned} \tag{36}$$

For every positive c_0 , we can find C' such that the following estimate holds for the linear term in (36):

$$\begin{aligned} & \left| \int (2\pi i r_{ss} \langle \xi_{ss}, v_{ss}(\gamma) \rangle + 2\pi i r_a \langle \xi_a, v_a(\gamma) \rangle) e(\langle \xi'_o - \xi_o, v_o(\gamma) \rangle) d\mu_1(\gamma) \right| \\ & \leq \left| \int 2\pi i r_{ss} \langle \xi_{ss}, v_{ss}(\gamma) \rangle + 2\pi i r_a \langle \xi_a, v_a(\gamma) \rangle d\mu_1(\gamma) \right| + C(r_{ss} + r_a) |\xi'_o - \xi_o| \\ & \leq C'(r_{ss} + r_a)^2 + c_0 |\xi'_o - \xi_o|^2. \end{aligned} \tag{37}$$

For the second inequality, we used (C) to show that the first term vanishes, and the inequality between the geometric and arithmetic means to estimate the second term.

Consider the function

$$\Phi(\xi) = \int e(\langle \xi, v_o(\gamma) \rangle) d\mu_1(\gamma)$$

on V_o . Note that Φ depends only on μ_1 . Combining (36) and (37) we get

$$\|T_{r_{ss}, r_a, \xi'_o} \varphi\|_2 \leq \Phi(\xi'_o - \xi_o) \varphi + C'(r_{ss}^2 + r_a^2) + c_0 |\xi'_o - \xi_o|^2.$$

If $\Phi(\xi) = 1$ for some $\xi \neq 0$, then $\langle \xi, v_o(\gamma) \rangle$ is an integer for all $\gamma \in \text{supp}(\mu_1)$ which contradicts to (35), hence impossible. Using Taylor series expansion, (C) and the moment condition, we can show that $\Phi(\xi) \leq 1 - c_1 |\xi|^2$ for some $c_1 > 0$ and $|\xi| < R$. These estimates prove the lemma if we set $c_0 < c_1$. \square

In the following lemma our argument will be based on compactness and continuity therefore we have to restrict its scope to a finite dimensional subspace.

Denote by $\mathcal{P}_\alpha \subset L^2(S^{ss} \times S^a)$ the space consisting of functions $\varphi(\xi_{ss}, \xi_a)$ which are restrictions of polynomials P of degree at most α defined on $V_{ss} \oplus V_a$. By abuse of notation, we will also write \mathcal{P}_α for the (infinite dimensional) subspaces of \mathcal{H}_1 and \mathcal{H}_2 consisting of those functions which belong to \mathcal{P}_α as a function on $S^{ss} \times S^a$ for every fixed value of ξ_o . In addition, we denote by π_α the orthogonal projection to \mathcal{P}_α .

Let $X_\alpha \subset V_o$ be the set of those ξ_o for which there is $\varphi \in \mathcal{P}_\alpha$ such that $T_{0,0,\xi_o} \varphi = \varphi$. If $T_{0,0,\xi_o} \varphi = \varphi$ and $T_{0,0,\xi'_o} \varphi' = \varphi'$ for $\xi_o \neq \xi'_o$, then φ and φ' are both eigenfunctions of $T_{0,0,\xi_o}$ with different eigenvalues hence they are orthogonal. Since \mathcal{P}_α is finite dimensional, X_α is finite. Moreover, it is seen easily that $X_0 = \{0\}$.

Write $X_\alpha = Y_\alpha \cup Z_\alpha$ where Y_α is the largest subset of X_α which is invariant under $\theta_o(\text{supp}(\mu))$.

Recall the definitions of δ and \mathcal{H}_2 from the beginning of the section. In the next Lemma and throughout below, we assume that $C^{-1} > \delta > C(r_{ss} + r_a)$ for any fixed constant $C > 1$ which may depend on μ, α and R .

Lemma 34. *There is $c > 0$ which depends only on μ and R such that the following hold. Let $\varphi \in \mathcal{H}_2$ with $\|\varphi\|_{\mathcal{H}_2} = 1$. Then there is $0 \leq l \leq |Z_\alpha|$ such that*

$$\|S_{r_{ss}, r_a} \pi_\alpha S_{r_{ss}, r_a}^l \varphi\|_{\mathcal{H}_2} \leq 1 - c\delta^2.$$

Proof. It is clear that there is some $c_\alpha > 0$ such that for each $\xi_o \in Z_\alpha$, there is $l \leq |Z_\alpha|$ such that there is $\gamma \in \text{supp}(\mu^{*(l)})$ with $\text{dist}(\theta_o(\gamma)^{-1}\xi_o, X_\alpha) > c_\alpha$. Using this, it is easy to show that there is $0 \leq l \leq |Z_\alpha|$ such that

$$\int_{S^{ss} \times S^a} \int_{\text{dist}(\xi_o, Z_\alpha) > c_\alpha/2} |S_{r_{ss}, r_a}^l \varphi(\xi_{ss}, \xi_a, \xi_o)|^2 d\xi_{ss} d\xi_a d\xi_o > \varepsilon^2$$

for some $\varepsilon > 0$ depending on μ and α . Indeed, if the inequality is not already true for $l = 0$, then a $1/(2|Z_\alpha|)$ proportion of the L^2 mass of φ is concentrated near a point $\xi_o \in Z_\alpha$ and this L^2 mass is moved away from X_α by $\theta_o(\gamma)^{-1}$ with some positive $d\mu^{*(l)}(\gamma)$ probability.

Write $\psi = \pi_\alpha S_{r_{ss}, r_a}^l \varphi$. It will be enough to show that

$$\|T_{\mathcal{H}_2} \psi\|_{\mathcal{H}_2} \leq 1 - c\delta^2.$$

Where we denoted by $T_{\mathcal{H}_2}$ the operator acting on \mathcal{H}_2 which is defined by the same formula as S_{r_{ss}, r_a} but using the measure μ_1 instead of μ . (This notation is not used elsewhere.) Indeed, the lemma follows by the familiar argument turning an estimate for $T_{\mathcal{H}_2}$ into an estimate for S_{r_{ss}, r_a} .

By assumption, $\delta < c_\alpha/2$ hence an ε proportion of the L^2 -mass of ψ is δ away from X_α . Then it is enough to show that

$$\|T_{r_{ss}, r_a, \xi_o} \psi'\|_2 \leq 1 - c\delta^2$$

for any $\psi' \in \mathcal{P}_\alpha \subset L^2(S^{ss} \times S^a)$ with $\|\psi'\|_2 = 1$ and $\text{dist}(\xi_o, X_\alpha) > \delta$. Indeed, the previous claim then follows by integrating ξ_o .

Let ξ'_o be the closest point of X_α to ξ_o . Write W for the 1-eigenspace of $T_{0,0,\xi'_o}$ in \mathcal{P}_α , and write U for the orthogonal complement. Write π_W and π_U for the orthogonal projections respectively. Set $a = \|\pi_W \psi'\|_2$ and $b = \|\pi_U \psi'\|_2$.

Since W and U are invariant under $T_{0,0,\xi_o}$, we have

$$\pi_U T_{0,0,\xi_o} \pi_W \psi' = 0 = \pi_W T_{0,0,\xi_o} \pi_U \psi'.$$

Consider the points of V_o to which ξ'_o is the closest among the points of X_α . In this region $T_{0,0,\xi_o}$ has norm less than 1 on the space U . A simple

continuity argument shows that there is a constant c_1 depending only on μ_1, α and R (independent of δ) such that

$$\|T_{0,0,\xi_o}\pi_U\psi'\|_2 < (1 - c_1)b.$$

Combining the above inequalities with Lemma 32 we get

$$\|\pi_U T_{r_{ss}, r_a, \xi_o} \pi_W \psi'\|_2 \leq C(r_{ss} + r_a)a \quad (38)$$

$$\|\pi_W T_{r_{ss}, r_a, \xi_o} \pi_U \psi'\|_2 \leq C(r_{ss} + r_a)b \quad (39)$$

$$\|\pi_U T_{r_{ss}, r_a, \xi_o} \pi_U \psi'\|_2 < (1 - c_1/2)b. \quad (40)$$

if r_{ss} and r_a are sufficiently small (depending on c_1).

From Lemma 33 we get

$$\|\pi_W T_{r_{ss}, r_a, \xi_o} \pi_W \psi'\|_2 < (1 - c\delta^2)a. \quad (41)$$

Combining estimates (38–41) we can write

$$\begin{aligned} \|T_{r_{ss}, r_a, \xi_o} \psi'\|_2^2 &\leq [(1 - c\delta^2)a + C(r_{ss} + r_a)b]^2 \\ &\quad + [(1 - c_1/2)b + C(r_{ss} + r_a)a]^2 \\ &\leq (1 - c\delta^2)a^2 + (1 - c_1/2)b^2 \\ &\quad + 4C(r_{ss} + r_a)ab + C^2(r_{ss} + r_a)^2 \\ &\leq (1 - c\delta^2/2)a^2 + (1 - c_1/2 + \frac{C_2(r_{ss} + r_a)^2}{\delta^2})b^2 \\ &\quad + C^2(r_{ss} + r_a)^2. \end{aligned}$$

In the last line, we used the inequality between the geometric and the arithmetic means. By the assumption (34) on δ , we have $10C_2(r_{ss} + r_a)^2 < c_1\delta^2$, and the lemma follows. \square

The last ingredient needed for the proofs of Propositions 30 and 31 is the following lemma which allows us to approximate the Fourier transform of $\mu^{*(l)}$ by polynomials in the ξ_{ss} and ξ_a variables.

Lemma 35. *There is a constant C depending on α and μ such that*

$$\int |v(\gamma)|^\alpha d\mu^{*(l)}(\gamma) \leq Cl^{\alpha/2}.$$

Proof. Let X_1, \dots, X_l be independent random isometries with law μ . Consider the sequence of random vectors

$$Y_l = v(X_1 \cdots X_l) = v(X_1) + \theta(X_1)v(X_2) + \dots + \theta(X_1) \cdots \theta(X_{l-1})v(X_l).$$

By (C), these form a martingale, and its conditional moments of order α are uniformly bounded. Thus the lemma follows from Burkholder's inequality, see [8, Theorem 3.2]. \square

Proof of Proposition 30. Let $\psi'_0 \in L^2(S^o)$ be the restriction of $\widehat{\delta}_{x_0}$ to the sphere of radius r_o in V_o , i.e.

$$\psi'_0(\xi_o) = e(\langle x_0, r_o \xi_o \rangle) = \widehat{\delta_{x_0}}(r_o \xi_o).$$

Let $S_{0,0,r_o}$ and $T_{0,0,r_o}$ be operators on $L^2(S^o)$ defined by

$$S_{0,0,r_o} \varphi(\xi_o) = \int e(\langle r_o \xi_o, v_o(\gamma) \rangle) \varphi(\theta_o(\gamma)^{-1} \xi_o) d\mu(\gamma)$$

and $T_{0,0,r_o}$ by a similar formula using μ_1 instead of μ . Then

$$(\mu^{*(l)} \cdot \delta_{x_0})^\wedge(r_o \xi_o) = S_{0,0,r_o}^l \psi'_0(\xi_o).$$

By a simple continuity argument similar to that in Lemma 33, we get a constant c depending on μ_1 and R such that $\|T_{0,0,r_o}\| < 1 - c$ for $R^{-1} < r_o < R$. Then $\|S_{0,0,r_o}\| < 1 - c$ follows with a perhaps worse constant.

Iterating the norm estimate for $S_{0,0,r_o}$, we get $\|S_{0,0,r_o}^{l_0} \psi'_0\|_2 \leq e^{-cl}$. Take $l_0 = d \log(s^{-1})/c$. Then

$$\|S_{0,0,r_o}^{l_0} \psi'_0\|_2 < s^d.$$

By Lemma 35, we have that

$$\|(\mu^{*(l_0)} \cdot \delta_{x_0})^\wedge\|_{\text{Lip}} < C l_0^{1/2}.$$

We can write

$$s^{2d} > \|S_{0,0,r_o}^{l_0} \psi'_0\|_2^2 > c_0 \|S_{0,0,r_o}^{l_0} \psi'_0\|_\infty^2 \cdot \left(\frac{\|S_{0,0,r_o}^{l_0} \psi'_0\|_\infty}{\|S_{0,0,r_o}^{l_0} \psi'_0\|_{\text{Lip}}} \right)^{\dim(V_o)-1},$$

where c_0 is a constant depending on the dimension. Thus

$$\|S_{0,0,r_o}^{l_0} \psi'_0\|_\infty < C s.$$

Using the Lipschitz norm estimate again, we get

$$|(\mu^{*(l)} \cdot \delta_{x_0})^\wedge(r_{ss}\xi_{ss} + r_a\xi_a + r_o\xi_o)| < Cs + C(r_{ss} + r_a)l_0^{1/2}.$$

We get the result after integrating the square of this over the domain

$$S^{ss} \times S^a \times \{\xi_o : R^{-1} \leq |\xi_o| \leq R, \text{dist}(\xi_o, Y_\alpha) \leq \delta\}$$

which is of measure $C\delta^{\dim(V_o)}$.

□

Proof of Proposition 31. By Lemma 35 and Taylor's theorem, we can write

$$S_{r_{ss}, r_a}^l \psi_2 = p_l + \omega_l,$$

such that $p_l \in \mathcal{P}_{\alpha-1} \subset \mathcal{H}_2$ and

$$\|\omega_l\|_2 < Cl^{\alpha/2}(r_{ss}^\alpha + r_a^\alpha).$$

We define recursively the increasing sequence of integers l_i . Let $l_0 = 0$ and if l_i is defined, then we put $l_{i+1} \leq l_i + |Z_\alpha| + 1$ such that by Lemma 34 we have

$$\|S_{r_{ss}, r_a}^{l_{i+1}} \psi_2\|_2 < (1 - c\delta^2)\|S_{r_{ss}, r_a}^{l_i} \psi_2\|_2 + \|\omega_{l_{i+1}-1}\|_2.$$

Combining the above, we get

$$\|S_{r_{ss}, r_a}^l \psi_2\|_2 \leq (1 - c\delta^2)^{l/(|Z_\alpha|+1)} + \sum_{j < l} \|\omega_j\|_2. \quad (42)$$

Take

$$l = C_1 \log(s^{-1})\delta^{-2}$$

for a suitably large constant C_1 depending on μ, α and R such that the first term in 42 is less than s^α . The second term is at most

$$Cl^{\alpha/2+1}(r_{ss}^\alpha + r_a^\alpha)$$

which was to be proved.

□

7.3 Proof of the Local Limit Theorem

Recall from the statement of the theorem that X_1, \dots are independent identically distributed random isometries. By the assumptions of the Theorem, the common law of X_i is non-degenerate and has finite moments of order $\alpha > d^2 + 3d$.

By Lemma 20, we can choose the origin in such a way that $v := \mathbf{E}[X_1(x_0) - x_0]$ is fixed by K . Now let $\gamma_v \in \text{Isom}(\mathbf{R}^d)$ be translation by $-v$ and consider the random isometries $X_i \cdot \gamma_v$ and denote by μ their common law. Then μ also satisfies (C) besides non-degeneracy and the above moment condition, and clearly it is enough to prove the theorem for these modified random isometries.

We can approximate any compactly supported continuous function in L^∞ norm by functions which have smooth (say C^∞) and compactly supported Fourier transform. Therefore we consider an arbitrary function f such that \widehat{f} is smooth and compactly supported, and prove the conclusion of Theorem 2 for it. Then this will prove the theorem by approximation. Let $R > 0$ be a number such that the support of \widehat{f} is contained in the ball of radius R around the origin.

We again write $\nu_l = \mu^{*(l)} \cdot \delta_{x_0}$ and use Plancherel's formula

$$\int f(x) d\nu_l(x) = \int \widehat{f}(\xi) \widehat{\nu}_l(\xi) d\xi.$$

Let Δ be the quadratic form from Proposition 10. It is easily seen that

$$\lim_{l \rightarrow \infty} l^{d/2} \int \widehat{f}(\xi) e^{-l\Delta(\xi, \xi)} d\xi = c\widehat{f}(0),$$

where c is a constant depending on Δ . Since $\widehat{f}(0) = \int f(x) dx$, it is enough to show that

$$\lim_{l \rightarrow \infty} l^{d/2} \int \widehat{f}(\xi) (\widehat{\nu}_l(\xi) - e^{-l\Delta(\xi, \xi)}) d\xi = 0.$$

The rest of the proof is devoted to estimating the above integral. We break it up into several regions. Let $\delta = l^{-\beta}$ with $\beta > d/(2d+2)$ and also

$$\beta(\alpha + 2) - \frac{\alpha}{2} < -\frac{d}{2},$$

which is possible since $\alpha > d^2 + 3d$. (This will also be the δ that we set in Propositions 30 and 31.) The first region is defined by

$$\Omega_1 := \{\xi : |\xi| \leq \delta\}.$$

Proposition 10 implies that

$$r^{-d+1} \int_{|\xi|=r} |\widehat{\nu}_l(\xi) - e^{-l\Delta(\xi,\xi)}|^2 d\xi \leq Cr^2.$$

By the Cauchy-Schwartz inequality, we have

$$r^{-d+1} \int_{|\xi|=r} |\widehat{\nu}_l(\xi) - e^{-l\Delta(\xi,\xi)}| d\xi \leq Cr.$$

After integrating for $0 \leq r \leq \delta = l^{-\beta}$ and using $|\widehat{f}(\xi)| \leq \|f\|_1$, we get

$$|\int_{\Omega_1} \widehat{f}(\xi)(\widehat{\nu}(\xi) - e^{-l\Delta(\xi,\xi)}) d\xi| \leq C\|f\|_1 l^{-\beta(d+1)}.$$

Since $\beta > d/(2d+2)$, the right side is $o(l^{-d/2})$.

Recall the notation from the beginning of Section 7, where we decomposed \mathbf{R}^d as an orthogonal sum $V_{ss} \oplus V_a \oplus V_o$. Moreover, we write $\xi = (\xi_{ss}, \xi_a, \xi_o)$, where ξ_i is the component of ξ in the corresponding subspace V_i .

The second region we consider is

$$\Omega_2 := \{\xi = (\xi_{ss}, \xi_a, \xi_o) : |\xi_{ss}| + |\xi_a| > C_0 l^{-1/2} \log^{1/2} l, |\xi| < R\},$$

where C_0 is a suitable constant depending on μ . (This region has an overlap with the first one.) Similarly to the proof of Lemma 23, we can deduce from Proposition 24 that there is a constant $c_0 > 0$ depending on μ and R such that

$$\begin{aligned} r_{ss}^{-(\dim V_{ss}-1)} r_a^{-(\dim V_a-1)} r_o^{-(\dim V_o-1)} \int_{|\xi_{ss}|=r_{ss}, |\xi_a|=r_a, |\xi_o|=r_o} |\widehat{\nu}_l(\xi)|^2 d\xi \\ \leq C e^{-c_0 \min\{(r_{ss}+r_a)^2 l, l^{1/4}\}} \end{aligned}$$

for $r_{ss}, r_a, r_o < R$. Integrating the above inequality and using the Cauchy-Schwartz inequality as above, we get

$$|\int_{\Omega_2} \widehat{f}(\xi)(\widehat{\nu}(\xi) - e^{-l\Delta(\xi,\xi)}) d\xi| \leq \|f\|_1 l^{-A},$$

where A can be any positive constant if we take C_0 sufficiently large. Note that $e^{-l\Delta(\xi,\xi)}$ is negligible in the region of integration.

Let X_α be the finite subset of V_o that we defined in Section 7.2. The third region we consider is given by the inequalities

$$\Omega_3 = \{\xi : |\xi_{ss}| + |\xi_a| < C_0 l^{-1/2} \log^{1/2} l, \text{ dist}(\xi_o, X_\alpha \setminus \{0\}) < \delta, |\xi| < R\}.$$

We apply Proposition 30 and get

$$r_{ss}^{-(\dim V_{ss}-1)} r_a^{-(\dim V_a-1)} \delta^{-\dim V_o} \int_{\xi \in \Omega_3, |\xi_{ss}|=r_{ss}, |\xi_a|=r_a} |\widehat{\nu}_l(\xi)| d\xi \leq C l^{-1/2} \log^{3/2} l$$

for any $r_{ss} + r_a < C_0 l^{-1/2} \log^{1/2} l$. (We also used the Cauchy Schwartz inequality as above.)

As in the previous two cases, we conclude

$$\begin{aligned} & \left| \int_{\Omega_3} \widehat{f}(\xi) (\widehat{\nu}(\xi) - e^{-l\Delta(\xi, \xi)}) d\xi \right| \\ & \leq C \|f\|_1 l^{-(\dim V_{ss} + \dim V_a)/2} l^{-\beta \dim V_o} l^{-1/2} \log^{(3 + \dim V_{ss} + \dim V_a)/2} l. \end{aligned}$$

By $\beta > d/(2d+2)$, the right hand side is less than $o(l^{-d/2})$.

The last region we consider is given by the inequalities

$$\Omega_4 = \{\xi : |\xi_{ss}| + |\xi_a| < C_0 l^{-1/2} \log^{1/2} l, \text{ dist}(\xi_o, X_\alpha) > \delta, |\xi| < R\}.$$

We apply Proposition 31 and the Cauchy Schwartz inequality, and get

$$r_{ss}^{-(\dim V_{ss}-1)} r_a^{-(\dim V_a-1)} \int_{\xi \in \Omega_4, |\xi_{ss}|=r_{ss}, |\xi_a|=r_a} |\widehat{\nu}_l(\xi)| d\xi \leq C \delta^{-\alpha-2} l^{-\alpha/2} \log^{\alpha+1} l.$$

If we integrate the above inequality similarly as above, we get

$$\left| \int_{\Omega_4} \widehat{f}(\xi) (\widehat{\nu}(\xi) - e^{-l\Delta(\xi, \xi)}) d\xi \right| \leq C \|f\|_1 o(l^{-d/2})$$

since

$$\beta(\alpha+2) - \frac{\alpha}{2} < -\frac{d}{2}.$$

Combining the estimates for the four regions above, we get the theorem.

References

- [1] M. Ådahl, I. Melbourne, M. Nicol, *Random iteration of Euclidean isometries*, Nonlinearity **16** (2003), no. 3, 977–987. [2](#)
- [2] V. I. Arnold, A. L. Krylov, *Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain* Dokl. Akad. Nauk SSSR **148** (1963) 9-12. [1](#)
- [3] P. Baldi, P. Bougerol, P. Crépel, *Théorème central limite local sur les extensions compactes de \mathbf{R}^d* , Ann. Inst. H. Poincaré Sect. B (N.S.) **14** (1978), no. 1, 99–111. [2](#)
- [4] H. Bass, *Theorems of Jordan and Burnside for algebraic groups*, J. Algebra **82** (1983), No. 1, 245–254. [5](#), [36](#)
- [5] E. Breuillard, *Random walks on Lie groups*, <http://www.math.u-psud.fr/~breuilla/part0gb.pdf> [9](#)
- [6] J. Bourgain, A. Gamburd, *On the spectral gap for finitely-generated subgroups of $SU(2)$* , Invent. math. **171** (2008), no. 1, 83-121. [12](#)
- [7] J. Bourgain, A. Gamburd, *A Spectral Gap Theorem in $SU(d)$* , <http://arxiv.org/abs/1108.6264> [12](#)
- [8] D. L. Burkholder, *Distribution Function Inequalities for Martingales*, Ann. Probab. **1**, (1973), no.1, 19–42. [5](#), [56](#)
- [9] J. P. Conze, Y. Guivarc’h *Ergodicity of group actions and spectral gap, applications to random walks and Markov shifts*, preprint, <http://arxiv.org/abs/1106.3248v1> [2](#), [12](#)
- [10] C. M. Dawson, M. A. Nielsen, *The Solovay-Kitaev algorithm*, Quantum Inf. Comput., **6**, No. 1, 81–95. (2006) [11](#)
- [11] D. Dolgopyat, *On mixing properties of compact group extensions of hyperbolic systems*, Israel J. Math. **130** (2002), 157–205. [11](#)
- [12] M. Gotô, *A theorem on compact semi-simple groups*, J. Math. Soc. Japan **1**, (1949) 270-272. [5](#), [45](#)

- [13] Y. Guivarc'h *Equirépartition dans les espaces homogènes*, Théorie ergodique (Actes Journées Ergodiques, Rennes, 1973/1974), pp. 131-142. Lecture Notes in Math., Vol. **532**, Springer, Berlin, 1976. [2](#), [9](#), [10](#), [46](#)
- [14] L. G. Gorostiza, *The Central Limit Theorem for Random Motions of d -Dimensional Euclidean Space*, Ann. Probability, **1**, 603–612 (1973) [2](#)
- [15] A. Grincevicius, *On the domain of the normal attraction of a stable law in the group of motions of a Euclidean space*, Liet. Matem. Rink. **25**, 39–52 (1985) [2](#)
- [16] D. Každan, *Uniform distribution on a plane*, Trudy Moskov. Mat. Obšč. **14** (1965) 299-305. [2](#), [9](#)
- [17] Yu. S. Khokhlov, *The domain of normal attraction of a semistable distribution on a semidirect product compact group and \mathbf{R}^d* , J. Math. Sci. **76** (1995), no. 1, 2147–2152. [2](#)
- [18] Yu. S. Khokhlov, *A local limit theorem for the composition of random motions of Euclidean space* Dokl. Akad. Nauk SSSR **260** (1981), no. 2, 295–299. English translation: Soviet Math. Dokl. **24**, no. 2, 277–281 (1982). [2](#)
- [19] A. Y. Kitaev, *Quantum computations: algorithms and error correction*. Russ. Math. Surv., **52** No. 6, 1191–1249 (1997) [11](#)
- [20] V. M. Maximov, *Local theorems for Euclidean motions. I* Z. Wahrsch. Verw. Gebiete **51** (1980), no. 1, 27–38. [2](#)
- [21] S. T. Rachev and J. E. Yukich, *Rates of Convergence of α -Stable Random Motions* J. Theoret. Probab. **4** (1991), no. 2, 333-352. [2](#)
- [22] B. Roynette, *Théorème central-limite pour le groupe des déplacements de \mathbf{R}^d* , Ann. Inst. H. Poincaré Sect. B (N.S.) **10** (1974), 391–398 [2](#)
- [23] V. N. Tutubalin, *The central limit theorem for random motions of Euclidean space*, Vestn. Mosk. Univ., Ser. 1 **22**, 100-108 (1967). [2](#), [10](#), [19](#)
- [24] P.P. Varjú, *Random walks in compact groups*, preprint <http://arxiv.org/abs/1209.1745> [11](#), [12](#)

CENTRE FOR MATHEMATICAL SCIENCES, WILBERFORCE ROAD, CAM-
BRIDGE CB3 0WA, ENGLAND

AND

THE EINSTEIN INSTITUTE OF MATHEMATICS, EDMOND J. SAFRA CAM-
PUS, GIVAT RAM, THE HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM,
91904, ISRAEL

e-mail address: pv270@dpmms.cam.ac.uk